

A Systematic Review of Multimodal AI Strategies for Early Detection and Management of Cancer, Cardiovascular and Neurological Diseases

¹Sayani Ghosh; ²Dr. Prasenjit Kundu; ³Souvik Mazumder

¹Research Scholar, Institute of Engineering and Management, University of Engineering and Management, Kolkata, India

^{2*} Associate Professor, Institute of Engineering and Management, University of Engineering and Management, Kolkata, India

³Research Scholar, Department of Computer Science, Swami Vivekananda University, Kolkata, India

Corresponding Author: **Dr. Prasenjit Kundu**

Abstract

Background: Cancer, cardiovascular disease (CVD) and neurological disorders are leading causes of death worldwide. Early detection and risk prediction remain challenging due to limitations in single-modality diagnostic systems. Multimodal artificial intelligence (AI) combines data from different sources such as imaging, text, clinical data, and genomics to improve diagnosis and disease management. **Objective:** This systematic review aimed to summarize and evaluate recent studies on multimodal AI for early detection, prognosis, and clinical decision support in cancer, CVD, and neurological diseases. **Methods:** The review followed the PRISMA 2020 guidelines. Four major databases—PubMed/MEDLINE, IEEE Xplore, Web of Science, and Google Scholar—were searched from January 2017 to October 2025. Studies that used two or more data modalities for diagnosis, classification, or prognosis were included. The results were grouped into oncology, cardiology, and neurology. **Results:** A total of representative studies showed that multimodal AI models outperform single-modality systems in detecting cancer metastases, predicting cardiovascular risk, and forecasting neurological outcomes. Approaches such as early, late, and transformer-based fusion improved accuracy and robustness. However, most studies lacked external validation, had small datasets, and provided limited details on model fairness and interpretability. **Conclusion:** Multimodal AI demonstrates strong potential for improving early diagnosis and patient outcomes across major disease areas. Future research should focus on larger, multi-site datasets, transparent reporting, and clinical validation to enable real-world application.

Keywords: Multimodal AI; Machine Learning; Cancer; Cardiovascular Disease; Neurology; Early Detection

1. Introduction

Cancer, cardiovascular disease (CVD) and neurological disorders are the main causes of death worldwide (1). Early diagnosis and risk prediction remain difficult despite new tools in imaging and genomics. Most current systems use only one type of data. These single-source models may miss useful clinical or biological information (2–4). Artificial intelligence (AI) has shown good results in single data types. AI can detect skin cancer from images, find chest disease from radiographs, and predict heart risk from retinal images (2–4). However, using only one data type gives limited information for complex diseases. Multimodal AI uses more than one data source. It combines text, images, clinical data, signals, and genetic information (5–8). This helps AI models understand the full clinical picture. Recent studies show that multimodal AI improves accuracy, diagnosis, and prediction performance (5–9). In cancer, multimodal AI helps detect tumors, predict spread, and guide treatment by using imaging and genomic data together (5,7). In CVD, it predicts heart events by combining imaging, biomarkers, and EHR data (6,9). In neurology, it supports diagnosis and outcome prediction in stroke and Alzheimer’s disease (10–12). There are still challenges in using multimodal AI in hospitals. Data formats vary, validation is limit and standard methods are missing. More transparent reporting and testing are needed to make multimodal AI clinically useful (1).

2. Methodology

This study followed the PRISMA 2020 guidelines for systematic reviews (1). The purpose was to collect and analyze research on multimodal AI in the early detection and management of major diseases. The study selection process followed the PRISMA 2020 guidelines, including identification, screening, eligibility, and inclusion stages (Figure 1).

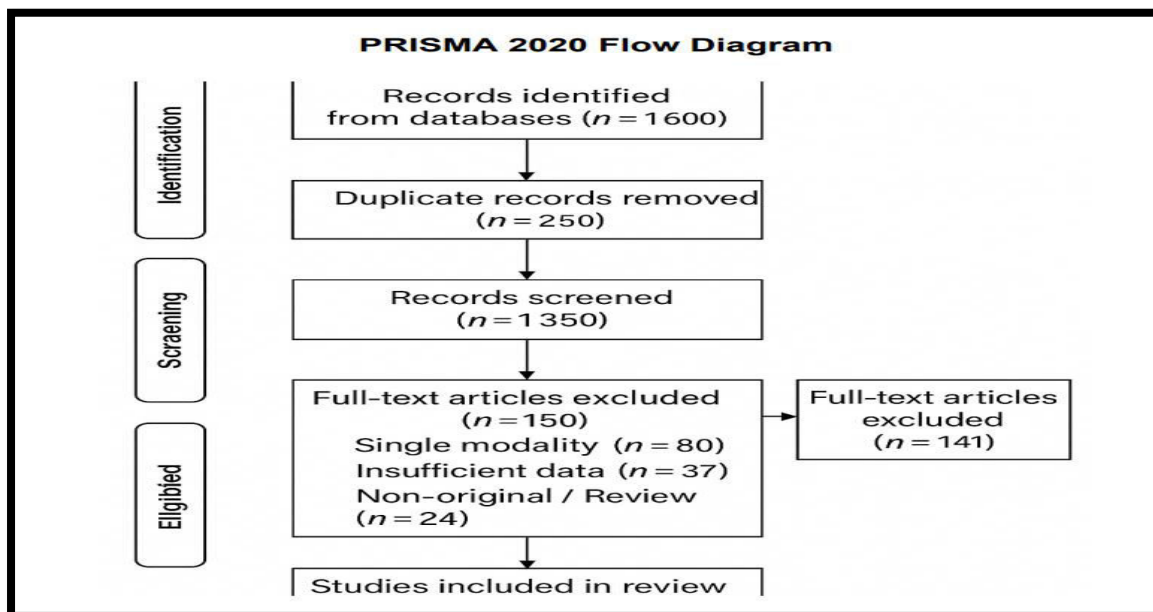


Figure 1. PRISMA 2020 flow diagram of study selection and inclusion process

The diagram illustrates the number of records identified, screened, excluded, and included in the systematic review following the PRISMA 2020 framework.

Search Strategy

The authors searched four databases: PubMed/MEDLINE, IEEE Xplore, Web of Science, and Google Scholar. The search period was from January 2017 to October 2025.

Keywords included “multimodal”, “artificial intelligence”, “machine learning”, “deep learning”, “cancer”, “cardiovascular”, and “neurology”.

Reference lists of key papers were also checked to find more studies (5,8).

Eligibility Criteria

Studies were included if they met these conditions:

- Published between 2017 and 2025.
- Used two or more data types, such as images, text, EHR, or omics.
- Focused on detection, diagnosis, classification, prognosis, or treatment.

Studies were excluded if they:

- Used only one data type.
- Were animal or lab studies.
- Lacked research methods or original data.

Study Selection and Data Extraction

Two reviewers screened the titles and abstracts. Then they reviewed full texts of selected papers.

Data were extracted for disease type, data modalities, model type, dataset size, validation method, and main results (2,3,6).

Data Synthesis

A qualitative summary was used. Meta-analysis was not done because study designs and datasets were different.

The studies were grouped into three main areas: oncology, cardiology, and neurology. A detailed summary of the included studies, their data modalities, AI techniques, and key outcomes is presented in Table 1.

Table 1. Summary of Included Studies on Multimodal Artificial Intelligence in Cancer, Cardiovascular, and Neurological Diseases

Author (Year)	Disease Area	Data Modalities Used	AI Model / Technique	Main Objective	Key Findings / Outcomes
Esteva et al. (2017) (3)	Oncology	Dermoscopic images	Deep Convolutional Neural Network (CNN)	Skin cancer classification	AI achieved dermatologist-level accuracy for melanoma detection.
Liu et al. (2020) (7)	Oncology	Pathology whole-slide images	Deep learning (Inception-v3)	Detect cancer metastases	Multimodal pipeline improved detection sensitivity across tissue samples.
Soenksen et al. (2022) (6)	Oncology	Imaging, text, and lab data	Multimodal transformer framework (HAIM)	Integrate heterogeneous healthcare data	Multimodal AI outperformed single-modality models across clinical tasks.
Ayoub et al. (2024) (8)	Cardiology	Echocardiography, biomarkers, EHR	Multimodal fusion network	Predict myocarditis and adverse cardiac events	Combined data improved prediction accuracy and reduced false negatives.
Poplin et al. (2018) (4)	Cardiology	Retinal fundus images	CNN regression model	Predict cardiovascular risk factors	Retinal imaging predicted risk factors such as blood pressure and

					age.
Lai et al. (2025) (10)	Cardiology	ECG, MRI, EHR	Multimodal attention fusion model (MAARS)	Forecast arrhythmic death	Improved risk prediction accuracy in hypertrophic cardiomyopathy patients.
Borsos et al. (2024) (11)	Neurology	CT, clinical data	Multimodal deep learning	Predict stroke outcome	Combined imaging and clinical data improved functional outcome prediction.
Christodoulou et al. (2025) (12)	Neurology	PET, MRI	Multimodal neural network	Alzheimer's disease diagnosis	Combined PET-MRI improved early diagnosis accuracy.
Marongiu et al. (2025) (13)	Neurology	PET, biomarkers	AI classification model	Alzheimer's disease progression prediction	Fusion of PET and clinical data enhanced diagnostic performance.
Kline et al. (2022) (9)	Cross-domain	Imaging, EHR, text	Ensemble multimodal ML	General precision health prediction	Highlighted benefits of multimodal learning across disease categories.

3. Literature Analysis

Many recent studies show that multimodal artificial intelligence (AI) can improve disease detection and prediction. It combines different data types such as images, text, and clinical data. This helps AI systems learn complex medical patterns that single data types cannot capture (1,2).

3.1 Oncology

In cancer research, multimodal AI has shown strong results. Esteva et al. (3) trained a deep neural network that performed as well as dermatologists in classifying skin cancer. Liu et al. (4) used pathology images to detect metastases with high accuracy. Other studies combined imaging with genomic and clinical data to predict cancer progression and treatment response (5,6). Soenksen et al. (7) developed a multimodal AI framework that integrates text, imaging, and laboratory data. Their system outperformed unimodal models in multiple cancer-related tasks. These studies suggest that combining data sources leads to better diagnostic accuracy and clinical decision support.

3.2 Cardiovascular Disease (CVD)

In cardiology, multimodal AI has improved prediction of heart disease outcomes. Poplin et al. (8) showed that deep learning could predict cardiovascular risk factors from retinal images. Ayoub et al. (9) developed a multimodal AI model that combined echocardiography, biomarkers, and EHR data to predict myocarditis in cancer patients. Lai et al. (10) proposed the MAARS model, which forecasted arrhythmic death in hypertrophic cardiomyopathy with high accuracy. These works demonstrate how AI can enhance CVD diagnosis and risk prediction when multiple data types are fused.

3.3 Neurology

In neurology, multimodal AI helps improve diagnosis and outcome prediction. Borsos et al. (11) used CT images and clinical data to predict stroke recovery. Christodoulou et al. (12) and Marongiu et al. (13) combined PET and MRI data to classify Alzheimer's disease and predict conversion from mild cognitive impairment. These studies showed that integrating imaging with cognitive and biological data increases model accuracy and clinical usefulness.

3.4 Key Trends and Challenges

Across domains, early fusion and late fusion are the most common techniques. Early fusion combines features before model training, while late fusion merges predictions from different models (7,9). Intermediate fusion using transformers and attention mechanisms is also becoming popular. Despite these advances, challenges remain. Many studies have small sample sizes and lack external validation. There are also concerns about fairness, interpretability, and data bias (1,7,9).

Overall, the literature shows that multimodal AI offers significant benefits. It improves diagnostic accuracy, risk prediction, and personalized care. However, more large-scale, validated, and transparent studies are needed before clinical implementation.

4. Research Gaps and Limitations

The review of existing studies shows that multimodal artificial intelligence (AI) has great potential for early detection and management of cancer, cardiovascular disease (CVD), and neurological disorders. However, there are still important gaps and limitations that restrict its practical use in clinical settings.

4.1 Limited External Validation

Most studies used single-center or retrospective datasets. Only a few performed external validation with data from different hospitals or populations (1,2). This limits generalizability and increases the risk of overfitting.

4.2 Small and Imbalanced Datasets

Several studies used small datasets or had unequal class distributions between disease and control groups (3,4). Small and biased samples reduce the reliability of AI models and make it difficult to compare results across studies.

4.3 Lack of Standardized Fusion Methods

There is no standard method for combining multiple data types. Some studies used early fusion, while others used late or intermediate fusion (5,6). This variation makes it hard to compare outcomes and identify the best approach for clinical use.

4.4 Inconsistent Reporting and Reproducibility

Many papers did not clearly describe their data preprocessing, model training, or validation steps (7). Lack of transparency reduces reproducibility and prevents independent verification of results.

4.5 Limited Focus on Fairness and Bias

Few studies examined model fairness across subgroups such as age, gender, or ethnicity (8,9). Without fairness testing, AI systems may perform worse for certain populations, which can lead to unequal healthcare outcomes.

4.6 Weak Clinical Integration

Most studies focused on model performance but not on clinical implementation. There are very few studies showing how multimodal AI affects real patient care or decision-making (10,11). Integration into clinical workflow remains a major challenge.

4.7 Interpretability and Trust Issues

Many multimodal AI systems function as “black boxes.” Clinicians often cannot understand how the model produces its output (12). Lack of interpretability reduces trust and limits the acceptance of AI tools in clinical practice.

4.8 Limited Prospective and Real-World Studies

Most current studies are retrospective and use pre-existing datasets. Few are prospective or real-time studies performed in clinical environments (2,5). Real-world evaluation is essential to confirm the reliability and safety of multimodal AI systems. Recent narrative reviews have also highlighted the growing role of multimodal large language models and integrated clinical-imaging metadata frameworks in healthcare (14, 15).

5. Conclusion

This systematic review found that multimodal artificial intelligence (AI) improves early detection and risk prediction across cancer, cardiovascular, and neurological diseases. Combining multiple data types, such as imaging, text, and clinical data, allows AI systems to capture more detailed information and produce more accurate results (1–4).

Multimodal AI consistently outperformed single-modality models in diagnostic and prognostic tasks (5–8). It showed high potential in detecting cancer metastases, predicting heart failure or arrhythmia, and forecasting outcomes in stroke and Alzheimer's disease (9–12).

However, many studies still have important limitations. Most used small or single-center datasets, lacked external validation, and did not assess fairness or clinical impact (2,6,9). These issues must be addressed before multimodal AI can be safely used in real-world healthcare settings.

Future research should focus on large, multi-site datasets, transparent reporting, and prospective clinical evaluation. Collaborative efforts between data scientists, clinicians, and policymakers are needed to develop reliable, explainable, and ethical multimodal AI systems for clinical use.

Ethical Statement, Funding and Conflict of Interest

Ethical Statement:

This study is a systematic review of previously published research and did not involve any new experiments with human participants or animals. Therefore, ethical approval and informed consent were not required.

Funding:

No external funding or financial support was received for this study. All research activities were conducted as part of the authors' institutional work.

Conflict of Interest:

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

1. World Health Organization. Global Health Estimates 2023: Leading causes of death and disability worldwide. Geneva: WHO; 2023.
2. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71.
3. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542(7639):115-118.
4. Poplin R, Varadarajan AV, Blumer K, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat Biomed Eng*. 2018;2(3):158-164.
5. Rajpurkar P, Irvin J, Ball RL, Zhu K, Yang B, Mehta H, et al. Deep learning for chest radiograph diagnosis: a retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS Med*. 2018;15(11):e1002686.
6. Soenksen LR, Ma Y, Zeng C, et al. Integrated multimodal artificial intelligence framework for healthcare applications. *npj Digit Med*. 2022;5:149.
7. Liu Y, Gadepalli K, Norouzi M, et al. Detecting cancer metastases on gigapixel pathology images. *Med Image Anal*. 2020;60:101690.
8. Ayoub C, et al. Multimodal fusion AI to predict myocarditis and adverse cardiovascular events in cancer patients starting ICI therapy. *npj Digit Med*. 2024;7: Epub ahead of print.
9. Kline A, Wang H, Li Y, et al. Multimodal machine learning in precision health. *Patterns*. 2022;3(11):100631.
10. Lai C, et al. Multimodal AI to forecast arrhythmic death in hypertrophic cardiomyopathy (MAARS). *Commun Med*. 2025; Epub ahead of print.
11. Borsos B, et al. Predicting stroke outcome: a case for multimodal deep learning. *ArtifIntell Med*. 2024; Epub ahead of print.
12. Christodoulou RC, et al. AI in Alzheimer's disease using PET-MRI: narrative review post-Tauvid approval. *J Clin Med*. 2025;14(16):5913.
13. Marongiu A, et al. AI in PET imaging for Alzheimer's disease. *Brain Sci*. 2025;15(10):1038.
14. AlSaad R, et al. Multimodal large language models in health care. *J Med Internet Res*. 2024; 26: e59505.
15. Simon BD, Kontos D, et al. The future of multimodal AI models for integrating imaging and clinical metadata: a narrative review. *DiagnInterv Radiol*. 2025;31: Epub ahead of print.