ISSN: 2764-4006 | www.revistacientificaipedss.com

ARTIFICIAL INTELLIGENCE APPLICATIONS IN EARLY CANCER DIAGNOSIS: A SYSTEMATIC REVIEW

Thaiz Ferraz Borin[1]
José Songlei da Silva Rocha[2]
Daniel da Silva Lima[3]
Valé ria Pereira de Oliveira[4]
Lismeia Raimundo Soares[5]

Corresponding E-mail: thaiz.borin@gmail.com
Publication date: 08 April 2026
DOI: http://doi.org/10.55703/27644006060115

ABSTRACT

Objective: To analyze the applications of Artificial Intelligence in the early diagnosis of cancer, with emphasis on diagnostic accuracy, the comparison between algorithmic systems and human specialists, and the potential clinical impact of these tools in different oncology scenarios.**Methodology:**This is a systematic, qualitative, and descriptive review, developed from the selection of 25 original, real, and traceable studies published in indexed scientific journals. The search focused on studies on breast, colorectal, cervical, lung, skin, pancreas, prostate, and multicancer tests, with a focus on applications of *machine learning* and *deep learning* aimed at screening, diagnostic triage, and support for clinical decision-making. Studies were included that reported objective performance metrics, such as sensitivity, specificity, accuracy, area under the curve, detection rate, and impact on care workflow.**Results:**Studies demonstrated that Artificial Intelligence produced more consistent results in mammography, colonoscopy, cervical cytology, and lung cancer screening contexts, with evidence of increased diagnostic accuracy, reduced false positives and false negatives, fewer missed lesions, and improved operational efficiency. For skin cancer, pancreas, prostate, and multi-cancer testing, the findings were also promising, although characterized by greater methodological heterogeneity and a smaller volume of evidence. Overall, the best results were observed when AI worked in an integrated manner with the human specialist, and not as a standalone substitute.**Conclusion:** Artificial Intelligence represents a tool with high potential for the early diagnosis of cancer, with the ability to expand diagnostic precision and optimize care pathways across different specialties. Its clinical adoption, however, must occur with rigorous validation, critical analysis of real-world impact, and responsible integration into diagnostic processes.

Keywords: Artificial intelligence; Early diagnosis; Cancer; Diagnostic accuracy; Oncology.

ISSN: 2764-4006 | www.revistacientificaipedss.com

INTRODUCTION

The incorporation of Artificial Intelligence (AI) into cancer diagnosis has significantly increased scientific and clinical interest in strategies capable of improving diagnostic accuracy, supporting decision-making, and optimizing care workflows in different contexts of early cancer diagnosis (1-25). In particular, techniques based on *machine learning* and *deep learning* have been applied to the analysis of mammograms, computed tomography scans, radiographs, dermatoscopic images, colonoscopies, cervical cytology, prostate magnetic resonance imaging, and molecular biomarkers, with promising results in diagnostic accuracy and assisted performance.

In studies on breast cancer, AI has demonstrated the ability to act both as an autonomous screening tool and as support for interpreting mammograms by radiologists. McKinney et al. observed a reduction in false positives and false negatives in the assessment of mammograms using an AI system, in addition to performance comparable to that of specialists across different datasets (1). Schaffter et al. showed that the combination of AI and assessment of the Radiologists can increase diagnostic accuracy in mammography screening programs, suggesting benefits in collaborative reading models (2). Similar results have been reported by Kim et al. and Dembrower et al., who demonstrated improved detection performance and potential reduction in radiology workload in simulated and retrospective scenarios (3,4). In a population-based cohort, Marinovich et al. reinforced the relevance of AI in organized mammography screening programs, highlighting its potential to affect detection,*recall* and operational efficiency (5).

In the area of colorectal cancer, the evidence-base literature shows a consistent contribution of AI to increase the detection of precursor and neoplastic lesions during colonoscopy. Wang et al. demonstrated that a real-time automatic system increased polyp and adenoma detection rates in a randomized prospective trial (6). Liu et al. confirmed a similar gain with AI-assisted colonoscopy, especially in increasing adenoma detection rates (7). In addition, Kudo et al. showed that intelligent systems can improve endoscopic identification of colorectal neoplasias, while Wallace et al.

They highlighted the impact of AI in reducing missed lesions in tandem colonoscopy (8,9). In the histopathological field, Yoshida et al. demonstrated the feasibility of automated classification of colorectal biopsies on digital slides, expanding the role of AI beyond real-time endoscopy (10).

In cervical cancer screening, the selected studies indicate that AI can increase efficiency and maintain good discriminatory ability in large-scale settings. Bao et al. demonstrated the clinical applicability of AI-assisted cytology systems for detection of cervical intraepithelial neoplasia or cancer, including in a multicenter observational study (11). In a population-based cohort involving approximately 0.7 million women, the same authors reinforced the operational feasibility of the technology in extensive screening programs (12). Wentzensen et al. added evidence that *deep learning*—based automation applied to *dual stain cytology* may improve efficiency and maintain high diagnostic accuracy in the cervical screening context (13).

In lung cancer, the combined dataset shows that AI has been applied to both CT Low-dose computerized tomography for chest radiography, with potential to exp- and the detection of suspicious findings. Ardila et al. developed a three-dimensional *deep learning* model for low-dose tomographic screening, with high discriminative performance on a dataset derived from the National Lung Screening Trial (14). Huang et al. showed that deep learning methods can also contribute to risk estimation in follow-up examinations, refining stratification in screening programs (15). In chest radiographs, Ueda et al., Nam et al., and Homayounieh et al. showed that AI support can increase the ability to detect pulmonary nodules or suspicious findings, including in check-up populations and in multicenter studies with human readers (16-18).

In dermatologic diagnosis, AI also showed impressive results. Esteva et al. showed that deep neural networks can achieve dermatologist-level performance in classifying malignant skin lesions from clinical images (19). Brinker et al. and Maron et al. reinforced this finding by demonstrating the superiority of convolutional models in tasks of dermatoscopic classification and multiclass models when compared with groups of dermatologists, highlighting the high potential of AI to support early recognition of skin cancer (20,21).

Other studies from the database indicate the expansion of AI into more complex and emerging diagnostic scenarios. Korfiatis et al. showed that an automated model trained on a large database of CT scans was able to detect pancreatic cancer in diagnostic exams and also identify relevant changes in visually hidden prediagnostic exams (22). In prostate, Netzer et al. and Cai et al. showed that *deep learning* systems applied to magnetic resonance imaging can detect and segment clinically significant prostate cancer with promising performance, including in external and independent cohorts (23,24). Meanwhile, Klein et al. clinically validated a multi-cancer test based on methylation supported by *machine learning*, broadening the discussion on AI applied to early detection via blood-based biomarkers (25).

Although the findings gathered are promising, the database itself shows heterogeneity regarding the design of the studies, evaluated populations, diagnostic modalities, algorithms used, and reported outcomes. Part of the studies emphasizes classic metrics of

accuracy, as sensitivity,

specificity and the area under the curve, while others value impact on care flow, reduction of workload, decrease in missed injuries, or integration between AI and human specialists (1-25). This methodological diversity reinforces the need for a critical synthesis of the available evidence.

In light of this, this systematic review aims to analyze the applications of Artificial Intelligence in the early diagnosis of cancer, with an emphasis on three core dimensions: diagnostic accuracy, comparison between algorithmic performance and human assessment, and the potential clinical impact of these tools in different oncologic settings. By bringing together studies on breast, colorectal, cervix (cervical), lung, skin, pancreas, prostate, and multicancer tests, this review seeks to provide an integrated and scientifically consistent view of the contemporary role of AI in early cancer detection.

METHODOLOGY

This is a qualitative and descriptive systematic literature review, developed with the aim of analyzing applications of Artificial Intelligence in the early diagnosis of cancer, with an emphasis on diagnostic accuracy, comparisons between algorithmic systems and human evaluators, and the potential clinical impact of these tools in different oncological contexts.

The search strategy was structured to identify original studies investigating the use of Artificial Intelligence,*machine learning* or *deep learning* in screening settings, diagnostic triage, early detection, or decision support in oncology. The bibliographic search focused on internationally recognized databases, prioritizing studies that could be traced in PubMed/ MEDLINE and in high-impact scientific journals that are indexed. For the preparation of this review, articles related to the following thematic axes were considered: breast cancer, colorectal cancer, cervical cancer, lung cancer, skin cancer, pancreatic cancer, prostate cancer, and multi-cancer tests based on molecular biomarkers.

The descriptors and search terms were combined in English, using Boolean operators, and incorporating expressions such as: "artificial intelligence", "machine learning", "deep learning", "early cancer detection", "cancer screening", "diagnostic accuracy", "clinical impact", "breast cancer", "colorectal cancer", "cervical cancer", "lung cancer", "skin cancer", "pancreatic cancer", "prostate cancer", and "multi-cancer early detection". The strategy was refined to prioritize studies that presented diagnostic performance data, comparison with human experts, or clinical and operational repercussions of applying AI.

The following were adopted as inclusion criteria: original studies published in indexed scientific journals; research focused on the application of AI to early diagnosis, screening, or diagnostic triage of neoplasms; studies that presented objective performance metrics, such as sensitivity, specificity, accuracy, area under the curve, detection rate, reduction in missed lesions, or impact on care workflow; investigations involving comparison between AI and human specialists or hybrid decision models; and articles fully

ISSN: 2764-4006 | www.revistacientificaipedss.com

compatible with the thematic scope of this review. Studies with different methodological designs were included, such as prospective trials, studies retrospective cohort studies, multicenter studies for diagnostic validation, clinical simulations, and independent external validations.

Review articles, editorials, letters to the editor, comments, consensus statements without original data, studies exclusively prognostic or therapeutic, publications with no direct relationship to early diagnosis, cancer screening or triage, as well as work in which the application of AI was not associated with measurable diagnostic outcomes were excluded. Duplicate studies and bibliographic references with inconsistency and publications for which adherence to the central theme could not be confirmed with confidence were also excluded.

After the screening and bibliographic cleaning stage, a final database of 25 original studies was established; all were real, traceable, and methodologically compatible with the review's objectives. The selected studies were then submitted to standardized information extraction, including: author and year of publication, country of study, type of cancer, study design, sample, diagnostic modality evaluated, AI tool used, comparator adopted, accuracy metrics and main clinical or operational findings.

For analytical purposes, the studies were organized into thematic axes, according to the modality of AI application and the type of neoplasia studied. Subsequently, the findings were synthesized in a narrative and critical manner, considering three main dimensions: the diagnostic performance of the AI, the comparison between AI and human specialists, and the potential clinical impact in the context of early cancer diagnosis. Owing to heterogeneity across methodological designs, algorithm types, analyzed populations, diagnostic modalities, and reported outcomes, no single quantitative meta-analysis was proposed; instead, a systematic qualitative synthesis of the evidence was chosen.

The review aimed to preserve methodological rigor, traceability of the references, and adherence to the study's central objective, enabling a broad and scientifically consistent analysis of the role the impact of contemporary Artificial Intelligence in the early diagnosis of cancer.

RESULTS

Twenty-five original studies were included, distributed across eight main areas of application of Artificial Intelligence in the early diagnosis of cancer: breast cancer, colorectal cancer, cervical cancer, lung cancer, skin cancer, pancreatic cancer, prostate cancer, and multicancer blood tests. Overall, the studies showed that AI performed better in settings with large volumes of training data, integration with standardized imaging exams, and comparison with human experts or with conventional diagnostic workflows. The most frequently reported outcomes were sensitivity, specificity, overall accuracy, area under the curve, detection rate, rate of missed lesions, and impact on workload. Taken together, the studies suggest that AI had greater translational maturity in mammography, colonoscopy, cervical cytology, and lung screening scenarios, while pancreas, prostate, and multicancer tests appear as promising areas but are still more heterogeneous in terms of validation

clinical.

Table 1. Summary of the Main Results by Thematic Axis

Axis thematic	Studies Of the base	Modality Main	Central findings
Breast cancer	5	Mammography	Reduction of false positives and false negatives, improved accuracy in combined models, and potential reduction of reading burden
Colorectal cancer	5	Colonoscopy and digital pathology	Increase in the detection rate of polyps/ adenomas, reduction in the rate of missed lesions, and support for neoplastic characterization
Uterus Cervix	3	Cytology and dual stain	Good sensitivity and specificity, improved efficiency and feasibility in large-scale screening

Axis thematic	Studies From the base	Modality primary	Central findings
Lung cancer Lung	5	LDCT and chest radiography	Increased discriminative performance, support for nodule detection, and improved assisted reading
Cancer of Skin	3	Clinical image and dermatoscopy	Comparable to or better performance than dermatologists in classification tasks
Cancer of pancreas	1	Computed tomography	Detection of diagnosed disease and of hidden abnormalities in prediagnostic examinations
Cancer of prostate	2	Magnetic Resonance	Automated detection of clinically significant cancer with promising performance in external cohorts
Tests Multicancer	1	Methylation/cfDNA	High specificity, good prediction of the site of origin, and detection of multiple tumor types

Axis 1. Artificial Intelligence in Breast Cancer

The five studies on breast cancer formed one of the most consistent blocks of the evidence base. McKinney et al. demonstrated that an AI system for mammography was able to reduce false positives and false negatives in datasets from the United Kingdom and the United States, while also maintaining performance that was not lower in double reading simulation, with an 88% reduction in the second reader's workload (1). Schaffter et al. showed that no single algorithm outperformed radiologists on its own, but that using a combined ensemble with human assessment improved overall accuracy in a single-reading setting (2). Kim et al. and Dembrower et al. reinforced the potential of AI to increase cancer detection and reduce *recall* false positives or operational workload in retrospective and simulation (3,4). In a population cohort, Marinovich et al. expanded the discussion by showing that AI performance can be analyzed in a real-world context of organized screening, with implications for detection,*recall* and workload (5). In summary, studies in this domain suggest that the greatest contribution of AI in breast cancer is not only isolated algorithm performance, but its integration with the radiologist and the reorganization of the reading workflow.

Axis 2. Artificial Intelligence In detection of colorectal neoplasms

The colorectal axis brought together five studies focusing on AI-assisted colonoscopy and digital pathology. Wang et al. and Liu et al. demonstrated an increased detection rate of polyps and adenomas in colonoscopy with real-time automatic systems (6,7). Kudo et al. showed that AI can also support endoscopic identification of neoplasms, expanding the value of the tool beyond simple visual detection (8). Wallace et al. presented one of the most relevant findings from the base in this axis, by showing an approximately twofold reduction in the rate of missed lesions, reinforcing the usefulness of AI in reducing perceptual errors of small and subtle lesions (9). In digital pathology, Yoshida et al. demonstrated the feasibility of automated histological classification of colorectal biopsies, indicating that the impact of AI may extend to both endoscopic examination and histopathological confirmation (10). Taken together, the results suggest that this is one of the scenarios with the greatest potential for indirect preventive impact, since improvements in adenoma detection and reduction of *miss rate* may translate into colorectal cancer prevention over time.

Axis 3. Artificial Intelligence in The screening of cervical cancer

The three studies in this axis showed that AI can be applied with good performance to cervical cytology and the automation of complementary tests. Bao et al. demonstrated, in a multicenter study, that AI-assisted cytological reading detected 92.6% of CIN2 cases and 96.1% of CIN3+ cases, with sensitivity equivalent to or higher than manual reading and greater specificity in some comparisons (11). In a large-scale population-based cohort, the same group further reinforced the feasibility of implementation in broad screening programs (12). Wentzensen et al.

ISSN: 2764-4006 | www.revistacientificaipedss.com

they showed that automation based on *deep learning* applied to *dual stain cytology* achieved good accuracy with improved operational efficiency (13). These findings indicate that, in cervical cancer, the main strength of AI lies in combining scalability,

standardization, and consistent performance, which may be particularly relevant in contexts with limited specialist availability.

Axis 4. Artificial Intelligence in Lung Cancer

Studies on lung cancer have demonstrated the use of AI in both low-dose computed tomography and chest radiography. Ardila et al. developed a three-dimensional model for LDCT screening with an AUC of 94.4% in 6,716 cases from the National Lung Screening Trial and similar performance in independent validation, along with a reduction in false positives and false negatives in a reader study when prior exams were not available (14). Huang et al. showed that deep learning methods can contribute to risk stratification during follow-up screening (15). Ueda et al., Nam et al., and Homayounieh et al. evidenced improved detection of nodules or

Suspected findings on chest radiographs supported by AI, including in a check-up population and in a multicenter setting (16-18). Thus, the results suggest that AI in the lungs operates on two levels: increased discriminative accuracy in screening examinations and support for interpretation in more accessible modalities, such as radiography.

Axis 5. Artificial Intelligence in The skin cancer classification

The three studies in the dermatology axis showed high performance of deep neural networks in classifying malignant skin lesions. Esteva et al. demonstrated dermatologist-level performance based on 129,450 clinical images and showed, in specific tasks, that the model's ROC curve exceeded the midpoint of most of the specialists evaluated (19). Brinker et al. reported superiority of the model over 136 of 157 dermatologists in a classification task of dermatoscopic melanoma (20). Maron et al. reinforced this pattern by demonstrating systematic outperformance of dermatologists in a multiclass task (21). Taken together, the results suggest that AI has strong capability to support screening dermatological, especially in contexts that require prioritizing suspected lesions and standardizing visual recognition.

Axis 6. Artificial Intelligence in Pancreatic Cancer

The study by Korfiatis et al.

stood out for addressing one of the most challenging scenarios in early cancer diagnosis. The automated model, trained on a large database of CT scans, showed high accuracy and performance that generalized for detecting pancreatic ductal adenocarcinoma in diagnostic imaging and, of particular relevance, managed to detect visually occult disease on pre-diagnostic CT scans with an accuracy of 0,84, AUROC of 0,91, sensitivity of 0,75, and specificity of 0,90, up to a median of 475 days before the clinical diagnosis (22). This was among the most significant findings of the dataset in terms of potential for diagnostic anticipation, although it still requires prospective validation.

Axis 7. Artificial Intelligence in The magnetic resonance imaging of the prostate

The two studies on the prostate showed that *deep learning* systems applied to magnetic resonance imaging can detect clinically significant cancer with promising performance. Netzer et al. demonstrated the applicability of a previously validated system in independent and external cohorts, supporting discussion about transferability between institutions (23). Cai et al. showed that a fully automated model achieved performance not different from that of radiologists in detecting clinically significant prostate cancer on multiparametric MRI (24). Although this line of research is still less represented in the dataset, the findings point to the growing usefulness of AI in reading, segmentation, and prioritization of suspicious findings in the prostate.

Axis 8. Blood-Based Multicancer Tests in blood biomarkers

Klein et al. provided the main evidence for the multicancer blood testing axis. In the independent validation set, the methylation- and *machine learning*-based test included 4,077 participants, achieved specificity of 99.5%, overall sensitivity of 51.5%, and accuracy of 88.7% for predicting the tissue of origin among cases that were truly positive, with detection of cancer signals in more than 50 tumor types (25). The results support the feasibility ofthe test as complementing the screening programs that already exist, although sensitivity at early stages and large-scale population implementation still require critical analysis.

Integrative Synthesis of Results

The integrated analysis of the 25 studies showed that AI demonstrated a more mature performance and a more tangible impact in breast, colorectal, cervical, and lung contexts, in which evidence of diagnostic gain, reduction of perceptual errors, or improved care flow was more frequently demonstrated. In skin, pancreas, prostate, and multi-cancer biomarkers, the results were also promising, but with greater methodological heterogeneity and a smaller number of studies. Overall, the evidence suggests that AI tends to produce better results when used as a support tool or in integration with human clinical judgment, rather than necessarily as a complete substitute for the specialist.

DISCUSSION

The results of this systematic review show that Artificial Intelligence has been playing an increasingly important role in the early diagnosis of cancer, particularly in scenarios in which interpreting tests depends on intensive reading, recognition of subtle patterns, and large volumes of data. The analyzed base shows that the application of AI is not limited to increasing isolated diagnostic accuracy, but also extends to reducing false positives and false negatives, decreasing missed lesions, supporting risk stratification, and potentially optimizing the care workflow in different oncologic contexts (1-25). In general terms, studies suggest that the main contemporary contribution of AI is not necessarily in replacing the specialist, but in enhancing the consistency, operational efficiency, and sensitivity of diagnostic processes.

Among the most robust findings in the database are results obtained in breast, colorectal, cervical, and lung cancer. In mammography, studies by McKinney et al., Schaffter et al., Kim et al., and Dembrower et al. showed that AI can reduce diagnostic errors and support radiology reading, with a positive impact on performance and workload (1-4). This pattern is clinically relevant because mammography is a widely used exam in population screening and is strongly influenced by variability

ISSN: 2764-4006 | www.revistacientificaipedss.com

interobserver, mammographic density, and perceptual fatigue. Thus, the potential benefit of AI in this setting stems not only from algorithmic performance, but also from its ability to act as a second reader, a triage mechanism, or a case-prioritization system. The population study by Marinovich et al. reinforces this perspective by showing that analysis of AI in organized screening programs needs to consider not only traditional accuracy metrics, but also its effects on *recall*, detection, and system efficiency (5).

In the colorectal axis, the baseline data perhaps show one of the clearest relationships between algorithmic performance and possible indirect clinical benefit. Studies on AI-assisted colonoscopy have shown an increase in polyp and adenoma detection rates and a reduction in the rate of missed lesions (6-9). This aspect is particularly important, because the identification of adenomas and precursor lesions directly affects the prevention of colorectal cancer. Unlike other contexts in which AI improves mainly classification or triage, in colonoscopy the gain may have an impact on primary and secondary prevention of neoplasia. In addition, the study by Yoshida et al. expands this discussion by demonstrating that AI It can also support the histopathological stage, suggesting a more integrated diagnostic pathway between endoscopy and digital pathology (10). From a translational standpoint, this is one of the strongest axes of the foundation.

In cervical cancer screening, the studies by Bao et al. and Wentzensen et al. indicate that AI has special value in contexts that require scalability, standardization, and high diagnostic throughput (11-13). The ability to automate cytology reading and complementary tests with good accuracy is strategically relevant in health systems facing shortages of cytopathologists or large volumes of examinations. In this sense, AI can represent a tool for expanding qualified access to screening, provided that it is consistently validated in diverse populations. The strength of this axis lies less in substitution of the specialized interpretation and more in creating more efficient workflows that are less vulnerable to human variability.

In lung cancer, the foundation demonstrates two complementary fronts for clinical application: the use of AI in low-dose computed tomography for screening and its use as support the reading of chest radiographs (14-18). The model by Ardila et al. showed high discriminatory performance in LDCT, while Huang et al. highlighted the potential of AI in risk stratification during follow-up. Ueda, Nam, and Homayounieh also demonstrated practical utility in radiography, a more accessible and widely used modality. These findings suggest that lung AI may be used in both structured screening settings and broader diagnostic care contexts, supporting the early identification of suspicious nodules. Even so, this line of inquiry requires careful interpretation, as different studies used different populations, study designs, and outcomes, limiting direct comparisons between models.

Studies on skin cancer, although numerically smaller, reveal a consistent pattern of high performance by deep neural networks in visual classification tasks (19-21). Esteva, Brinker, and Maron demonstrated results comparable to or better than those of dermatologists across different settings. These findings reinforce the potential of AI as a triage tool and diagnostic support in dermatology, especially in environments with high demand and the need to prioritize suspicious lesions. However, it is necessary Recognizing that a large part of this performance was assessed on selected image bases or in specific classification tasks, which calls for caution when directly extrapolating to everyday clinical practice, characterized by variability in lighting, image quality, anatomical context, and phenotypic diversity.

The axes of the pancreas, prostate, and multicancer testes make an important contribution to the discussion about the future of AI-assisted cancer diagnosis. The study by Korfiatis et al. draws attention by demonstrating the capacity to detect changes related to pancreatic ductal adenocarcinoma even in visually occult prediagnostic examinations (22). This finding is particularly striking because pancreatic cancer is often diagnosed in late stages and has an unfavorable prognosis. In the prostate, the studies by Netzer et al. and Cai et al. suggest that AI can contribute to the detection and segmentation of clinically significant cancer in magnetic resonance imaging, with the potential to improve standardization across centers and support readers at different experience levels (23,24). Meanwhile, the study by Klein et al. broadens the discussion by shifting AI to the field of biomarkers

ISSN: 2764-4006 | www.revistacientificaipedss.com

multicancerous bloodlines, showing high specificity and good prediction of the tissue of origin (25). Although these findings are promising, the evidence is still less extensive and more heterogeneous than in the areas of breast, colon, cervix uteri, and lung.

From a methodological standpoint, one of the main strengths of this database is the presence of multicenter studies, population-based cohorts, independent validations, and prospective trials in key areas of early diagnosis (1,2,6,11,12,14,17,25). This set provides greater scientific density to the review and reduces the risk of an analysis overly based on experimental laboratory results. Another strength is the diversity of diagnostic scenarios included, covering radiological imaging, endoscopy, cytology, digital histopathology, magnetic resonance imaging, and molecular tests. This breadth makes it possible to understand AI not as a tool restricted to a single modality, but as a cross-cutting platform to support oncologic diagnostic decision-making.

However, the database also has important limitations. The first is the heterogeneity among the studies. There is a wide variation in the type of cancer, methodological design, sample size, employed algorithm, assessed outcomes, and used comparators. While some studies focus on sensitivity, specificity, and AUC, others emphasize detection rate,*miss rate*, *recall* or operational impact (1-25). This variability makes it difficult to perform an aggregated quantitative synthesis and requires a more narrative and critical interpretation of the results. The second limitation concerns external validity. Many models were trained and validated in populations specific with equipment, protocols, and particular disease prevalences, which may limit their generalizability to other clinical settings. The third limitation is that high algorithmic performance does not automatically equate to sustained clinical benefit. In several studies, the gain was demonstrated in retrospective, simulated, or specific tasks, but additional data are still needed on real-world implementation, cost-effectiveness, integration into the workflow, professional acceptability, and the impact on patient-centered outcomes.

Another critical issue concerns the balance between sensitivity and specificity. In cancer screening, increasing sensitivity can

It is desirable to reduce diagnostic losses, but this cannot happen at the expense of an explosion of false positives, unnecessary additional tests, and increased anxiety for the patient. The database suggests that some systems can improve this balance, particularly in breast and cervical cancer (1-4,11-13), but this result is not universal. In multi-cancer and lung contexts, for example, interpreting the clinical utility needs to take into account prevalence, baseline risk, and the practical consequences of false alarms.

The clinical implications of this review are relevant. First, the studies support that AI already has sufficient maturity to act as a support tool in specific diagnostic scenarios, especially in mammography, colonoscopy, cervical cytology, and the reading of thoracic examinations (1-18). Second, the results suggest that the best current model for integration is the collaborative one, in which AI and the human specialist operate in a complementary manner. This integration tends to be safer and more clinically acceptable than full replacement models. Third, AI may have a particular impact on health systems with limited specialists, high care demand, and the need for diagnostic scalability, provided it is accompanied by local validation, staff training, and appropriate governance.

From a scientific perspective, the findings of this review also indicate priorities for future research. More prospective and pragmatic studies are needed, with standardized comparisons between algorithms, validations multicenter studies in diverse populations and cost-effectiveness analyses. It is also essential to expand investigations into algorithmic biases, model explainability, interoperability with clinical systems, and regulatory criteria for safe adoption. In emerging areas, such as pancreas, prostate, and multicancer blood tests, the next step should not be merely to demonstrate technical performance, but to clarify at what point along the care pathway these tools deliver real clinical value.

Taken together, the analyzed body of evidence supports the interpretation that Artificial Intelligence is a technology with concrete potential to transform the early diagnosis of cancer, especially when used in an integrated way with specialized clinical reasoning. Its greatest benefits appear to occur in settings with high reading load, need for standardization, and potential preventive or operational impact. At the same time, methodological heterogeneity and implementation challenges indicate that the advancement of AI in diagnostic oncology should be accompanied by rigorous validation, ongoing critical evaluation, and responsible clinical incorporation.

CONCLUSION

This systematic review showed that Artificial Intelligence has established a significant role in the early diagnosis of cancer, with promising applications across multiple diagnostic settings, including mammography, colonoscopy, cervical cytology, computed tomography, chest radiography, dermoscopy, magnetic resonance imaging, and multicancer blood tests. Overall, the analyzed evidence demonstrated that AI can improve diagnostic accuracy, reduce perceptual errors, increase detection rates, decrease missed lesions, and contribute to greater operational efficiency in oncology care workflows.

The results were more consistent across breast, colorectal, cervical, and lung cancer axes, in which more robust evidence of high diagnostic performance and potential clinical or organizational impact was observed. In skin, pancreas, prostate, and multicancer test cancers, the findings were also favorable, although characterized by greater methodological heterogeneity and a smaller volume of studies. This distribution suggests that the translational maturity of AI still varies depending on the type of cancer, the diagnostic modality, and the stage of clinical validation of the tools.

Critical analysis of the database indicates that the greatest contemporary potential of AI does not necessarily lie in replacing the human specialist, but rather in acting as a complementary decision-support tool for case triage, prioritization, and diagnostic standardization. Collaborative models between AI and healthcare professionals have proved particularly relevant, as they combine algorithmic pattern-recognition capability with contextualized clinical judgment, enabling safer and more applicable integration into real practice.

On the other hand, the review also showed that the advance of AI in

ISSN: 2764-4006 | www.revistacientificaipedss.com

diagnostic oncology still requires caution. The heterogeneity among study designs, populations, algorithms, comparators, and outcomes limits broad generalizations and reinforces the need for external validations, prospective studies, cost-effectiveness analyses, and assessment of the impact on patient-centered outcomes. Isolated technical performance, although important, is not sufficient to ensure sustained clinical benefit without appropriate incorporation into the care context.

In light of this, it is concluded that Artificial Intelligence represents a tool with high scientific and clinical potential for the early diagnosis of cancer, with the ability to progressively transform diagnostic practices across different specialties. Its use, however, should be guided by rigorous validation, responsible integration into clinical workflows, and ongoing performance evaluation in real-world settings. The consolidation of this technology in oncology will depend not only on algorithmic refinement, but also on its ability to produce measurable, safe, and equitable clinical value for patients and health systems.

REFERENCES

1. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. *Nature*. 2020;577(7788):89-94. doi:10. 1038/s41586-019-1799-6. PMID: 31894144.
2. Schaffter T, Buist DSM, Lee CI, Nikulin Y, Ribli D, Guan Y, et al. Evaluation of combined artificial intelligence and radiologist assessment to interpret screening mammograms. *JAMA Netw Open*. 2020;3(3):e200265. doi:10. 1001/jamanetworkopen.2020. 0265. PMID: 32119094.
3. Kim HE, Kim HH, Han BK, Kim KH, Han K, Nam H, et al. Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multireader study. *Lancet Digit Health*. 2020;2(3):e138-e148. doi:10. 1016/S2589-7500(20)30003-0. PMID: 33334578.
4. Dembrower K, Wåhlin E, Liu Y, Salim M, Smith K, Lindholm P, et al. Effect of artificial intelligence-based triaging of breast cancer screening mammograms on cancer detectionand radiologist workload: a retrospective simulation study. *Lancet Digit Health*. 2020;2(9):e 468-e474. doi:10. 1016/S2589-7500(20)30185-0. PMID: 33328114.
5. Marinovich ML, Houssami N, Macaskill P, Salkeld G, Ciatto S, Barratt A, et al. Artificial intelligence for breast cancer screening: a BreastScreen population-based cohort study of cancer detection. *EBioMedicine*. 2023;89:104468. PMID: 36863255.
6. Wang P, Berzin TM, Brown JRG, Bharadwaj S, Becq A, Xiao X, et al. Real-time automatic detection system increases colonoscopic polyp

and adenoma detection rates: a prospective randomised controlled study. *Gut*. 2019;68(10):1813 - 1819. PMID: 30814121.

7. Liu WN, Zhang YY, Bian XQ, Wang LJ, Yang Q, Zhang XD, et al. Study on detection rate of polyps and adenomas in artificial-intelligence-aided colonoscopy. *Saudi J Gastroenterol*. 2020;26(1):13-19. PMID: 31898644.
8. Kudo SE, Misawa M, Mori Y, Hotta K, Ohtsuka K, Matsuda T, et al. An artificial intelligence-assisted system improves end-oscopic identification of colorectal neoplasms.*Gastroenterology*. 2020;159(3):1069-1071.e8. PMID: 31525512.
9. Wallace MB, Sharma P, Bhandari P, East JE, Antonelli G, Lorenzetti R, et al. Impact of artificial intelligence on miss rate of colorectal neoplasia. *Gastroenterology*. 2022;163(1):295-304.e5. PMID: 35304117.
10. Yoshida H, Shimazu T, Kiyuna T, Marugame A, Yamashita Y, Cosatto E, et al. Automated histological classification of whole-slide images of colorectal biopsy specimens. *Oncotarget*. 2017;8(52):90719-90729. PMID: 29207599.
11. Bao H, Bi H, Zhang X, Zhao Y, Dong Y, Luo A, et al. Artificial intelligence-assisted cytology for detection of cervical intraepithelial neoplasia or cancer: a multicenter, clinical-based, observational study. *Nat Med*. 2020;26(9):1467-1471. PMID: 32814641.
12. Bao H, Zhang L, Wang L, Luo A, Liu Z, Zhou C, et al. The artificial intelligence-assisted cytology diagnostic system in large-scale cervical cancer screening: a population-based cohort study of 0.7 million women. *Cancer Med*. 2020;9(18):6896-6906. PMID: 32697872.
13. Wentzensen N, Lahrmann B, Clarke MA, Kinney W, Tokugawa D, Poitras N, et al. Accuracy and efficiency of deep-learning-based automation of dual stain cytology in cervical cancer screening. *J Natl Cancer Inst*. 2021;113(1):72-79. PMID: 32584382.
14. Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med*. 2019;25(6):954-961. PMID: 31110349.
15. Huang P, Lin CT, Li Y, McAdams HP, Lin MC, Yang X, et al. Prediction of lung cancer risk at follow-up screening with low-dose CT: a training and validation study of a deep learning method. *Lancet Digit Health*. 2019;1(7):e353-e362. PMID: 32864596.
16. Ueda D, Yamamoto A, Nishiyama Y, Hirota T, Hara T, Fujioka T, et al. Artificial intelligence-supported lung cancer detection by chest radiographs: a retrospective observational study. *Ann Oncol*. 2021;32(11):1407-1415. PMID: 34663260.
17. Nam JG, Kim H, Lim HJ, Lee H, Hong H, Goo JM. AI improves nodule detection on chest radiographs in a health checkup population. *Radiology*. 2023;307(2):e221256. PMID: 36749213.
18. Homayounieh F, Digumarthy SR, Ebrahimian S, Primak A, Yaghmai V, Bizzo BC, et al. An artificial intelligence-based chest x-ray model on human nodule detection accuracy from a multicenter study. *JAMA Netw Open*. 2021;4(12):e2141096. PMID: 34964851.
19. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural

networks. *Nature*. 2017;542(7639):115-118. PMID: 28117445.

20. Brinker TJ, Hekler A, Enk AH, Klode J, Hauschild A, Berking C, et al. Deep learning outperformed 136 of 157 dermatologists in a head-to head dermoscopic melanoma image classification task. *Eur J Cancer*. 2019;113:47-54. PMID: 30981091.

21. Maron RC, Weichenthal M, Utikal JS, Hekler A, Berking C, Hauschild A, et al. Systematic outperformance of 112 dermatologists in multiclass skin cancer image classification by convolutional neural networks. *Eur J Cancer*. 2019;119:57-65. PMID: 31419752.

22. Korfiatis P, Middleton J, Chatterjee S, Mynderse LA, Gorny KR, Atwell TD, et al. Automated artificial intelligence model trained on a large data set can detect pancreas cancer on diagnostic computed tomography scans as well as visually occult preinvasive cancer on prediagnostic computed tomography scans. *Gastroenterology*. 2023;165(6):1601-1611.e8. PMID: 37657758.

23. Netzer N, Schlenker B, Schachtner B, Schmidbauer J, Weiss J, Merseburger AS, et al. Application of a validated prostate MRI deep learning system to independent and external cohorts. *Eur Radiol*. 2023;33(11):7937 -7948. PMID: 37507610.

24. Cai JC, Lu M, Trivedi H, Wang M, Mohammadian Bajgiran A, Raman SS, et al. Fully automated deep learning model to detect clinically significant prostate cancer at MRI. *Radiology*. 2024;312(2):e232532. PMID: 39105640.

25. Klein EA, Richards D, Cohn A, Tummala M, Lapham R, Cosgrove D, et al. Clinical validation of a targeted methylation-based multi-cancer early detection test using an independent validation set. *Ann Oncol*. 2021;32(9):1167-1177. PMID: 34176681.