Deep Learning-Based Multimodal Neuroimaging Analysis for Early Detection of Neurodegenerative Diseases Using Brain CT and Clinical Data

Ewelina Smoczyk
School of Information Technology, University of Cincinnati, Cincinnati, OH, USA.
helloewelina@uc.edu

Lintang Rahimah
Department of Computer Science, University of Alabama at Birmingham, Birmingham, AL, USA.
lintangwork@uab.edu

Kacper Bugiel
Department of Computer Science, University of Houston, Houston, TX, USA.
kacper.bugiel642@uh.edu

Abstract

Early detection of neurodegenerative diseases remains a difficult clinical and infrastructural challenge because the earliest biological changes often precede obvious symptoms and because access to advanced imaging is unevenly distributed. This paper examines deep learning-based multimodal neuroimaging analysis for early detection using brain computed tomography and clinical data as a system-level intervention. Although magnetic resonance imaging and positron emission tomography provide richer tissue contrast and molecular specificity, computed tomography is widely available, fast, and central to acute care pathways, making it a pragmatic anchor for scalable screening and triage. The proposed analytical perspective treats deep learning as one component within a socio-technical architecture that includes data acquisition, preprocessing, multimodal fusion, clinical decision support, governance, deployment, and long-term monitoring. The paper discusses architectural trade-offs among centralized cloud inference, edge deployment, and hybrid systems, with attention to latency, privacy, bandwidth, maintenance, and resilience. It also analyzes fairness and robustness risks arising from site-specific imaging protocols, demographic skew, missing clinical variables, and distributional drift across hospitals. Governance mechanisms such as prospective validation, human oversight, auditability, post-market surveillance, and transparent reporting are presented as necessary complements to model performance. The conclusion argues that successful early detection systems must be designed not merely as high-performing classifiers but as sustainable clinical infrastructures that align technical capability with reimbursement, workforce capacity, regulatory pathways, and equitable access. Such systems can improve triage and referral only if they are embedded responsibly within existing neurological and primary care workflows.

Keywords

multimodal neuroimaging; deep learning; brain CT; clinical data; neurodegenerative diseases; early detection; clinical decision support; fairness; governance; health infrastructure.

1. Introduction

Neurodegenerative diseases such as Alzheimer's disease, Parkinson's disease, and related disorders impose growing burdens on aging populations, health systems, and caregivers. Their clinical trajectories are often gradual, and the most effective windows for intervention may occur before substantial irreversible neuronal loss. Early detection therefore carries substantial promise, yet it also exposes difficult tensions between sensitivity, specificity, cost, and accessibility. Deep learning has become a central approach in medical image analysis because it can learn hierarchical representations from large and heterogeneous datasets, reducing reliance on handcrafted features and enabling pattern discovery across imaging modalities [1]. In neuroimaging, these capabilities have been applied to structural and functional data, often with magnetic resonance imaging or positron emission tomography as the primary source [2]. However, those modalities are not universally available in routine care, especially in community hospitals, rural clinics, and lower-resource settings. Brain computed tomography, by contrast, is one of the most widely deployed neuroimaging tools in acute and general hospital environments.

The central argument of this paper is that deep learning-based multimodal analysis using brain computed tomography and clinical data should be understood as a system-level clinical infrastructure rather than as an isolated algorithmic innovation. The technical literature often emphasizes model architecture, benchmark accuracy, and modality fusion, but clinical value depends on how a system is deployed, maintained, governed, and integrated into workflows. Machine learning in medicine has shown potential to improve diagnosis, prognosis, and resource allocation, but realization of that potential requires attention to implementation science, regulation, and organizational change [9]. High-performance medicine emerges when computational advances converge with data availability, clinician trust, and operational readiness [10]. For neurodegenerative disease detection, this convergence is especially fragile because early signs are subtle, labels are uncertain, and the consequences of false reassurance or false alarm can be significant.

A further complication is that brain computed tomography is often treated as a rule-out modality rather than a definitive diagnostic instrument. It is excellent for detecting hemorrhage, mass effect, hydrocephalus, and large structural lesions, but it offers lower soft-tissue contrast than magnetic resonance imaging. Nevertheless, computed tomography is frequently the first neuroimaging study performed in emergency departments and general hospitals. If deep learning can extract additional predictive signal from routine computed tomography when combined with clinical variables such as cognitive scores, gait assessment, laboratory markers, medication history, and family history, then it may support earlier referral and more efficient use of advanced imaging. This prospect aligns with the biological and clinical frameworks that define neurodegenerative disease through multiple markers and longitudinal change rather than a single test [5]. The challenge is to design multimodal systems that are robust, fair, explainable enough for clinical use, and sustainable under real-world constraints.

2. Clinical and Technical Background

The clinical diagnosis of neurodegenerative disease has evolved from symptom-based classification toward biomarker-supported and multidimensional frameworks. For Alzheimer's disease, research criteria now incorporate amyloid, tau, and neurodegeneration markers, while clinical criteria continue to emphasize cognitive decline and functional impairment [5]. Structural imaging has long played a role in excluding other causes and in tracking atrophy patterns associated with disease progression [6]. International working

groups have refined diagnostic criteria to include biomarkers and to distinguish clinical syndromes from underlying pathology [7]. In Parkinson's disease and related disorders, clinical diagnostic criteria emphasize motor features, supportive findings, and exclusion of alternative diagnoses, with increasing interest in prodromal markers [8]. These developments create both an opportunity and a burden for computational systems: they provide richer labels for supervised learning, but they also reveal that diagnostic ground truth is often probabilistic, longitudinal, and specialty-dependent.

From a technical perspective, deep learning has transformed medical image analysis by enabling end-to-end learning from imaging data. Convolutional neural networks, recurrent architectures, attention mechanisms, and transformer-based models have been applied to segmentation, detection, classification, and prediction tasks [2]. Surveys of deep learning in medical imaging emphasize the importance of data quality, annotation, preprocessing, and validation, while also noting persistent challenges in generalization and interpretability [3]. Foundational texts on deep learning for medical image analysis describe an ecosystem in which imaging data, clinical variables, and domain knowledge must be integrated thoughtfully rather than simply concatenated [4]. Neuroimaging adds particular complexity because brain structure and function vary across age, sex, genetics, comorbidities, scanner type, and acquisition protocol.

Brain computed tomography occupies a distinctive position in this ecosystem. It is fast, relatively inexpensive, widely available, and compatible with acutely ill patients who cannot undergo long magnetic resonance imaging sessions. It is also routinely used in emergency settings for suspected stroke, trauma, and altered mental status. These advantages make it attractive for scalable early detection or triage. However, computed tomography has limitations in soft-tissue contrast, which can obscure subtle atrophy, white matter changes, or early neurodegenerative patterns. Deep learning can potentially compensate by learning texture, shape, and contextual features, but such learning requires large, diverse, and well-curated datasets. The required reference in this paper illustrates how deep learning can be applied to computed tomography scanning images for intelligent detection and analysis of brain tumors [15]. That work is not the focus of this paper, but it demonstrates the broader feasibility of computed tomography-based deep learning in neuro-oncology and motivates careful extension to neurodegenerative disease, where lesions are often diffuse rather than focal.

Clinical data are equally important in multimodal analysis. Cognitive test scores, functional questionnaires, motor assessments, laboratory results, medication histories, and demographic variables can provide complementary information that is not visible in imaging. They can also serve as proxies for disease severity, comorbidities, and access to care. When fused with imaging, clinical variables may improve discrimination between normal aging, mild cognitive impairment, and early dementia. Yet clinical data introduce their own challenges, including missingness, inconsistent coding, site-specific documentation practices, and temporal misalignment between image acquisition and clinical assessment. A robust multimodal system must therefore handle missing modalities, variable time intervals, and noisy labels without becoming brittle or overconfident.

3. Multimodal Learning Architecture and Data Integration

The architecture of a multimodal neuroimaging system can be conceptualized in layers. The acquisition layer includes computed tomography scanners, picture archiving and communication systems, electronic health records, and clinical assessment instruments. The

preprocessing layer performs skull stripping, registration, intensity normalization, and quality control, often with modality-specific adaptations. The representation layer learns imaging features and clinical features, either separately or jointly. The fusion layer combines these representations through early, intermediate, or late fusion strategies. The decision layer produces risk scores, diagnostic categories, or referral recommendations. The interface layer presents results to clinicians, patients, or care coordinators. The governance layer oversees validation, monitoring, access control, and accountability. Each layer introduces trade-offs that affect clinical usefulness.

Early fusion combines raw or minimally processed data before representation learning, which can capture cross-modal interactions but may be sensitive to missing data and scale differences. Intermediate fusion learns separate representations and then combines them at one or more hidden layers, offering flexibility and robustness. Late fusion combines independent predictions, which can be easier to implement and interpret but may miss subtle interactions. In practice, hybrid approaches are often used. For example, an imaging encoder may learn structural features from computed tomography, while a clinical encoder processes tabular variables such as age, cognitive score, and vascular risk factors. Attention mechanisms can weight modalities dynamically depending on data quality, so that a high-quality image and a complete clinical record contribute differently than a low-quality image with missing laboratory values. Such dynamic weighting is conceptually appealing, but it also complicates validation because model behavior may vary across sites and patient groups.

Data integration is not merely a technical task. It requires semantic alignment across institutions, because the same clinical concept may be recorded differently in different electronic health record systems. Imaging metadata such as slice thickness, reconstruction kernel, and contrast administration can influence model predictions. Harmonization methods can reduce site effects, but they may also remove biologically meaningful variation. Federated learning offers one approach to multi-institutional training without centralizing raw data, but it introduces challenges in communication efficiency, model aggregation, and governance. Transfer learning from large-scale imaging datasets can improve performance when local data are limited, but domain shift remains a risk. The broader health data ecosystem has been described as a foundation for machine learning, yet the gap between data availability and clinically actionable insight remains substantial [18]. In neurodegenerative disease, this gap is widened by long preclinical periods and the need for longitudinal follow-up.

A further architectural consideration is uncertainty representation. Early detection systems should not simply output a binary label; they should communicate confidence, ambiguity, and the need for further evaluation. This is especially important when computed tomography findings are subtle or when clinical data are incomplete. Calibration is essential because a model that is accurate on average may still be unreliable for individual patients. Ensemble methods, Bayesian approximations, and out-of-distribution detection can help, but they add computational cost and complexity. The system must also distinguish between detection of a probable neurodegenerative pattern and exclusion of acute pathology such as hemorrhage or mass effect. A clinically useful architecture may therefore be multi-task, simultaneously performing acute abnormality detection, atrophy assessment, and risk stratification. Such multi-task learning can improve efficiency, but it requires careful label design and loss weighting, which are governance as much as engineering decisions.

4. System-Level Trade-Offs in Deployment

Deployment choices shape the clinical, ethical, and economic performance of deep learning systems. Centralized cloud inference offers scalability, easier updates, and access to powerful computational resources. It also raises concerns about data transfer, privacy, latency, and dependence on network connectivity. Edge deployment, where inference occurs on local servers or devices within the hospital, can reduce latency and improve privacy, but it may limit model size and complicate updates. Hybrid architectures can balance these concerns by performing urgent triage locally and more complex analysis centrally when connectivity and consent allow. In acute care, latency matters because delayed results can delay treatment. In outpatient screening, throughput and cost may matter more than milliseconds. The appropriate architecture therefore depends on the clinical pathway and the resources of the institution.

A major system-level trade-off is between sensitivity and specificity. Early detection systems may be tuned to maximize sensitivity to avoid missing incipient disease, but this can generate false positives that burden specialists and cause patient anxiety. Conversely, high specificity may reduce false alarms but miss early cases. The optimal threshold depends on disease prevalence, referral capacity, and the availability of confirmatory tests. In settings with limited neurology services, a highly sensitive but less specific model could overwhelm the system. In settings with abundant specialist capacity, earlier and broader referral may be acceptable. This means that model thresholds should not be fixed globally; they should be adjustable based on local epidemiology and resource constraints. Such adaptation requires governance mechanisms that allow local calibration without fragmenting the evidence base.

Another trade-off involves integration with existing workflows. A system that requires radiologists to perform additional manual steps may be abandoned despite good accuracy. A system that integrates directly into the picture archiving and communication system or electronic health record may be more sustainable. However, deep integration raises cybersecurity and vendor lock-in concerns. Interoperability standards such as DICOM and HL7 Fast Healthcare Interoperability Resources can facilitate integration, but implementation quality varies. Clinical decision support should present results at the point of decision, with clear explanations and recommended next steps. It should also avoid automation bias, where clinicians accept model output uncritically. The literature on delivering clinical impact with artificial intelligence emphasizes that successful deployment depends on workflow fit, leadership, and continuous evaluation rather than algorithmic novelty alone [17]. Regulatory frameworks increasingly recognize the need for lifecycle oversight, as reflected in action plans for artificial intelligence and machine learning-based software as a medical device [14].

Sustainability is a further deployment concern. Computed tomography scanners, servers, storage systems, and network infrastructure consume energy and require maintenance. Large deep learning models can be computationally expensive to train and update. If every hospital deploys its own high-end inference infrastructure, costs may be duplicated and environmental impacts amplified. Shared regional inference services, federated learning, and model compression can reduce energy and cost, but they may introduce new dependencies. The system should also be resilient to scanner downtime, network outages, and cyberattacks. In disaster or rural settings, a cloud-only system may fail precisely when it is needed. A hybrid approach with local fallback capability is often more robust. These infrastructure choices are not peripheral; they determine whether a model can be sustained beyond a research pilot.

5. Governance, Fairness, and Robustness

Fairness is a central concern in multimodal neuroimaging because disparities can arise at multiple points. Historical inequities in access to care may mean that underrepresented groups

have less imaging data, fewer specialist evaluations, and less consistent documentation. Algorithms trained on such data can learn biased patterns and then reproduce or amplify them. A well-known example from population health management showed how an algorithm used health costs as a proxy for need and thereby disadvantaged Black patients [11]. In neuroimaging, similar risks exist if models rely on socioeconomic proxies, scanner access, or referral patterns. Fairness cannot be guaranteed by removing sensitive variables, because structural inequities may be encoded in imaging protocols, comorbidity patterns, and missing data. Instead, fairness must be evaluated across demographic groups, clinical sites, and disease subtypes using metrics that reflect both harms and benefits.

Robustness is closely related to fairness. A model may perform well on internal validation but fail when applied to a new scanner, a different patient population, or a changed clinical protocol. Distributional shift can occur gradually, as equipment ages or coding practices change, or abruptly, as during a pandemic or policy reform. Continuous monitoring is therefore necessary, including performance dashboards, drift detection, and periodic revalidation. Robustness also includes resilience to adversarial inputs, corrupted images, and missing clinical fields. In a clinical setting, adversarial attacks may be less likely than accidental data quality problems, but both can cause harm. Red teaming, stress testing, and simulation of rare scenarios can help identify failure modes before deployment.

Explainability is often proposed as a solution to trust and accountability, but it is not a panacea. Current explainable artificial intelligence methods can produce plausible visualizations that do not faithfully reflect model reasoning [13]. For clinical users, explanations should answer actionable questions: which findings drove the recommendation, how confident is the model, and what additional information would change the assessment. A multidisciplinary perspective on explainability in healthcare emphasizes that explanations must serve different audiences, including clinicians, patients, regulators, and developers [16]. Ethical challenges in implementing machine learning in health care include informed consent, transparency, liability, and equitable access [12]. These are not merely compliance issues; they shape whether patients and clinicians will trust and use the system.

Governance should therefore be designed as a lifecycle process. Before deployment, a system should undergo prospective validation in the intended clinical setting, with clear inclusion and exclusion criteria. During deployment, human oversight should be maintained, especially for high-stakes decisions. After deployment, post-market surveillance should track safety, effectiveness, and equity. Responsible machine learning roadmaps emphasize multidisciplinary collaboration, continuous evaluation, and attention to social context [20]. Ethical frameworks for machine learning in medicine call for accountability mechanisms that extend beyond technical teams [19]. For neurodegenerative disease, governance must also address the long-term implications of early diagnosis, including psychological impact, insurance discrimination, and the availability of interventions. Detecting risk without a pathway to care can create harm rather than benefit.

6. Infrastructure, Sustainability, and Policy Implications

The sustainability of deep learning-based early detection depends on infrastructure and policy as much as on model performance. Health systems need reliable imaging equipment, trained technologists, data storage, cybersecurity, and informatics support. They also need clinicians who can interpret and act on model outputs. In many regions, the shortage of neurologists and geriatricians is a limiting factor. A system that identifies more at-risk patients without expanding referral capacity may simply shift bottlenecks. Policy should therefore consider

workforce planning, telemedicine, and shared-care models. Reimbursement is another critical lever. If there is no billing pathway for computational analysis or risk assessment, hospitals may lack financial incentives to adopt and maintain such systems. Conversely, value-based payment models may reward earlier detection if it reduces costly crises and institutional care.

Data sharing policies also matter. Multi-institutional datasets improve generalizability, but they require privacy-preserving mechanisms, data use agreements, and community engagement. Federated learning and distributed analytics can enable collaboration without centralizing identifiable data, but they require technical standards and trust. Regulatory science must keep pace with adaptive algorithms that change over time. A static approval may be insufficient for models that are continuously updated. The U.S. Food and Drug Administration action plan on artificial intelligence and machine learning-based software as a medical device highlights the need for predetermined change control plans and lifecycle oversight [14]. Similar discussions are occurring in other jurisdictions. For neurodegenerative disease, regulatory pathways must also account for the fact that early detection may not immediately change treatment; it may instead inform monitoring, lifestyle intervention, or enrollment in research.

Environmental sustainability is an often overlooked dimension. Medical imaging contributes to greenhouse gas emissions through energy consumption and equipment manufacturing. Large-scale deep learning adds computational demand. If multimodal systems are deployed widely, their carbon footprint could be significant. Strategies such as model compression, efficient inference, shared regional data centers, and scheduling non-urgent analyses during low-carbon energy periods can reduce impact. Sustainability also means maintaining models over time. A model that is abandoned after a pilot wastes resources and erodes trust. Long-term funding, technical support, and governance must be planned from the beginning. The challenge is to create systems that are not only accurate but also maintainable, adaptable, and equitable over years of clinical use.

7. Case Illustrations and Cross-Domain Comparisons

Comparisons with other clinical domains can clarify the opportunities and limits of computed tomography-based multimodal analysis. In acute stroke, computed tomography is central to triage, and deep learning has been applied to hemorrhage detection, large vessel occlusion prediction, and early ischemic change assessment. These applications benefit from clear time-critical decisions and well-defined outcomes. In neuro-oncology, deep learning has been used for tumor detection and analysis from computed tomography scans, as illustrated by recent work in the field [15]. That case shows that computed tomography can support meaningful automated analysis when lesions are focal and training data are adequate. Neurodegenerative disease is different because pathology is diffuse, progression is slow, and clinical labels are often uncertain. A model may detect atrophy patterns, white matter changes, or vascular burden, but these findings overlap with normal aging and other comorbidities.

Cross-domain lessons also emerge from cardiovascular risk prediction and population health. Those fields have shown that multimodal data can improve risk stratification, but they have also revealed that algorithmic bias and poor calibration can undermine trust [11]. In oncology, artificial intelligence has achieved strong performance in narrow tasks, yet integration into clinical pathways remains uneven [17]. The gap between benchmark performance and clinical impact is a recurring theme. For neurodegenerative disease, this gap is likely to be wider because the disease is heterogeneous and because early detection may not lead to immediate disease-modifying treatment. The value proposition must therefore be framed carefully:

benefits may include better prognostic counseling, earlier planning, enrollment in clinical trials, and management of vascular risk factors.

A useful comparison can also be made with screening programs for other conditions. Screening is justified when the condition is serious, detectable early, treatable, and when the benefits outweigh harms. For neurodegenerative disease, treatability remains limited, although lifestyle and risk factor interventions may help. This means that early detection systems should not be marketed as definitive diagnostic tools. They should be positioned as risk stratification and triage aids that support clinical judgment. In primary care, where most patients first present, a computed tomography-based model combined with cognitive screening could help identify who needs specialist referral. In specialist settings, the same model could support longitudinal monitoring. The same technical system may serve different purposes, but each purpose requires different validation, thresholds, and governance.

8. Future Perspectives and Research Agenda

Future research should move beyond single-center retrospective studies toward prospective, multi-site, and pragmatic trials. Such trials should evaluate not only diagnostic accuracy but also clinical outcomes, workflow efficiency, cost-effectiveness, and equity. They should include diverse populations and a range of scanner types and protocols. They should also assess human factors, such as how clinicians interact with model outputs and whether automation bias or alert fatigue emerges. Explainability research should focus on clinically meaningful and faithful explanations rather than visually appealing heatmaps [13]. Responsible machine learning frameworks provide a starting point for organizing these efforts around transparency, accountability, and continuous improvement [20]. Deep learning in medical image analysis will continue to advance, but the field needs stronger evidence standards for clinical translation [3].

Another priority is the development of robust multimodal representations that can handle missing data and temporal irregularity. Neurodegenerative disease unfolds over years, so models should incorporate longitudinal trajectories rather than static snapshots. This may require sequential models, survival analysis, and joint modeling of imaging and clinical outcomes. Privacy-preserving learning across institutions will be essential for building diverse datasets. Synthetic data and simulation can augment rare cases, but they must be validated carefully to avoid introducing artifacts. Standardization of imaging protocols and clinical variables would improve comparability, but it must not stifle local innovation. A balance is needed between harmonization and site-specific adaptation.

Policy and governance research should also advance. There is a need for clear pathways for regulatory review of adaptive algorithms, reimbursement models for computational diagnostics, and liability frameworks when clinicians and algorithms disagree. Patient and public engagement is essential, especially for neurodegenerative disease, where early diagnosis carries personal and social consequences. International collaboration can help align standards, share best practices, and reduce duplication. Ultimately, the success of deep learning-based multimodal neuroimaging will depend on whether it can be embedded in sustainable, equitable, and trustworthy clinical infrastructures. Technical excellence is necessary but not sufficient. The field must design for governance, maintenance, and long-term value from the outset.

9. Conclusion

Deep learning-based multimodal neuroimaging analysis using brain computed tomography and clinical data offers a pragmatic pathway toward earlier detection and triage of neurodegenerative diseases. Computed tomography is widely available and integrated into acute and general care, while clinical data provide complementary information about cognition, function, and risk. When combined through carefully designed multimodal architectures, these data sources can potentially identify patients who need further evaluation, support referral decisions, and improve longitudinal monitoring. However, the paper has argued that such systems must be understood as socio-technical infrastructures rather than isolated models. Their value depends on deployment architecture, workflow integration, fairness, robustness, governance, sustainability, and policy alignment.

The most important trade-offs are not only technical but also organizational and ethical. Centralized and edge deployment each have advantages and limitations. Sensitivity and specificity must be calibrated to local prevalence and capacity. Explainability must serve real clinical decisions rather than provide false reassurance. Fairness must be evaluated across groups and sites, and robustness must be monitored continuously. Regulatory and reimbursement frameworks must support lifecycle oversight and sustainable adoption. Without these elements, high-performing models may fail to deliver clinical benefit or may even exacerbate disparities. With them, computed tomography-based multimodal deep learning could become a valuable component of a broader strategy for early neurodegenerative disease detection, provided that it is developed and governed responsibly.

References

1. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. https://doi.org/10.1038/nature14539

2. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (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

3. Shen, D., Wu, G., & Suk, H.-I. (2017). Deep learning in medical image analysis. Annual Review of Biomedical Engineering, 19, 221–248. https://doi.org/10.1146/annurev-bioeng-071516-044442

4. Zhou, S. K., Greenspan, H., & Shen, D. (Eds.). (2017). Deep learning for medical image analysis. Academic Press.

5. Jack, C. R., Jr., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., ... Sperling, R. (2018). NIA-AA research framework: Toward a biological definition of Alzheimer's disease. Alzheimer's & Dementia, 14(4), 535–562. https://doi.org/10.1016/j.jalz.2018.02.018

6. Frisoni, G. B., Fox, N. C., Jack, C. R., Jr., Scheltens, P., & Thompson, P. M. (2010). The clinical use of structural MRI in Alzheimer disease. Nature Reviews Neurology, 6(2), 67–77. https://doi.org/10.1038/nrneurol.2009.215

7. Dubois, B., Feldman, H. H., Jacova, C., Hampel, H., Molinuevo, J. L., Blennow, K., ... Cummings, J. L. (2014). Advancing research diagnostic criteria for Alzheimer's disease: The IWG-2 criteria. The Lancet Neurology, 13(6), 614–629. https://doi.org/10.1016/S1474-4422(14)70090-0

8. Postuma, R. B., Berg, D., Stern, M., Poewe, W., Olanow, C. W., Oertel, W., ... Deuschl, G. (2015). MDS clinical diagnostic criteria for Parkinson's disease. Movement Disorders, 30(12), 1591–1601. https://doi.org/10.1002/mds.26424

9. Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. New England Journal of Medicine, 380(14), 1347–1358. https://doi.org/10.1056/NEJMra1814259

10. Topol, E. J. (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

11. Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366(6464), 447–453. https://doi.org/10.1126/science.aax2342

12. Char, D. S., Shah, N. H., & Magnus, D. (2018). Implementing machine learning in health care—Addressing ethical challenges. New England Journal of Medicine, 378(11), 981–983. https://doi.org/10.1056/NEJMp1714229

13. Ghassemi, M., Oakden-Rayner, L., & Beam, A. L. (2021). The false hope of current approaches to explainable artificial intelligence in health care. The Lancet Digital Health, 3(11), e745–e750. https://doi.org/10.1016/S2589-7500(21)00208-9

14. U.S. Food and Drug Administration. (2021). Artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD) action plan. https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device

15. Hong, Ziwei, et al. "Intelligent Detection and Analysis of Brain Tumors Based on Deep Learning for CT Scanning Images." 2026 7th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT). IEEE, 2026.

16. Amann, J., Blasimme, A., Vayena, E., Frey, D., & Madai, V. I. (2020). Explainability for artificial intelligence in healthcare: A multidisciplinary perspective. BMC Medical Informatics and Decision Making, 20, 310. https://doi.org/10.1186/s12911-020-01332-6

17. Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. BMC Medicine, 17, 195. https://doi.org/10.1186/s12916-019-1426-2

18. Beam, A. L., & Kohane, I. S. (2018). Big data and machine learning in health care. JAMA, 319(13), 1317–1318. https://doi.org/10.1001/jama.2017.18391

19. Vayena, E., Blasimme, A., & Cohen, I. G. (2018). Machine learning in medicine: Addressing ethical challenges. PLoS Medicine, 15(11), e1002689. https://doi.org/10.1371/journal.pmed.1002689

20. Wiens, J., Saria, S., Sendak, M., Ghassemi, M., Liu, V. X., Doshi-Velez, F., ... Goldenberg, A. (2019). Do no harm: A roadmap for responsible machine learning for health care. Nature Medicine, 25(9), 1337–1340. https://doi.org/10.1038/s41591-019-0548-8