# Quantum Machine Learning for Alzheimer's Disease Diagnosis: A Critical Appraisal of Quantum-Advantage Claims

Manoj Dhiman
*Department of Computer Science and Informatics, Central University of Himachal Pradesh*
doi.org/10.64643/IJIRTV13I1-204007-459

***Abstract*—Quantum machine learning (QML) has been advanced as a route to faster and more accurate diagnosis of Alzheimer's disease (AD), and a rapidly growing body of primary studies reports near-perfect classification accuracies. This paper systematically identifies peer-reviewed primary studies applying genuine quantum methods to AD or mild cognitive impairment diagnosis and critically appraises the evidential basis of their quantum-advantage claims. Following PRISMA 2020, six databases were searched, yielding 865 records; fifteen studies met all criteria. Everyone was executed on a classical simulator, none reported unambiguous results on real quantum hardware, and only one modelled device noise. Datasets were frequently very small and ten of the eleven MRI-based studies used slice-level rather than subject-level splits, creating a tangible data-leakage risk. Seven studies reported accuracies at or above 98%, yet classical baselines were weak or effectively absent in a substantial minority, and only five of fifteen reported any measure of statistical uncertainty. The most methodologically careful study found an advantage smaller than its own error bars, and another reported near-chance performance. The AD-specific QML literature does not currently provide credible evidence of a quantum advantage; we propose a reporting standard to place the field on a sounder evidential footing.**

***Index Terms*—Alzheimer's disease, data leakage, neuroimaging, quantum machine learning, systematic review.**

## I. INTRODUCTION

Alzheimer's disease is a progressive neurodegenerative disorder and the most common cause of dementia. Early and accurate diagnosis is clinically valuable, and machine learning applied to neuroimaging and other biomarkers has consequently become a large and mature research field [3,17].

Quantum computing has more recently been proposed as a way to extend these methods. Quantum machine learning encompasses variational quantum classifiers, quantum kernel methods, quantum neural networks, and hybrid quantum-classical architectures [5]. The theoretical appeal is twofold: quantum feature maps can embed data into high-dimensional Hilbert spaces that may be hard to access classically [11], and under specific, well-characterised conditions certain quantum learning routines carry provable advantages [19]. These arguments have motivated a fast-growing stream of primary studies that apply QML to AD diagnosis, many reporting classification accuracies approaching or exceeding 99%.

Such results, taken at face value, would imply that quantum methods already outperform mature classical pipelines on a clinically important task. This claim warrants scrutiny for two reasons. First, contemporary quantum hardware is noisy and limited in qubit count, and most QML experiments are run on classical simulators; training variational circuits is itself difficult, with vanishing-gradient "barren plateau" effects a recognised obstacle [18]. Second, the AD neuroimaging field has well-documented methodological hazards — small cohorts, dataset overlap, and information leakage between training and test partitions — that can inflate apparent performance independently of the learning algorithm [16,27]. A recent scoping review found that such leakage produces a pervasive illusion of near-perfect performance in classical deep-learning studies of AD [4].

Existing reviews have not resolved this tension. Current syntheses of QML in neurology and healthcare are predominantly descriptive, and the single most rigorous critical review of QML in digital health deliberately excluded the neuroimaging-based AD literature [9]. This review addresses that gap: we provide the first systematic, AD-specific corpus of peer-reviewed QML primary studies under a transparent PRISMA 2020 protocol; we critically appraise each study on the dimensions that determine whether a quantum-advantage claim is credible; and we test whether the critical conclusion reached for digital-health QML replicates in the AD-specific setting.

## II. RELATED REVIEWS AND THE RESEARCH GAP

Surveys of QML in medical image analysis and of QML in healthcare provide useful taxonomies but treat reported accuracies largely as given, with AD appearing only incidentally [32,33]. Reviews of quantum computing in clinical care and of simulated quantum computing in neurological disease are broader in scope and moderately critical — noting the reliance on simulators — but are not AD-specific and do not audit baselines, partitioning, or uncertainty [8,14]. A systematic review of quantum deep learning in neuroinformatics is closest in subject matter but reports only a slight aggregate advantage and presents study-level accuracies without a study-by-study methodological critique [21].

Two reviews are strongly critical. A survey of QML in medicine and healthcare warns that headline accuracies of 95–99% should not be read as proof of quantum superiority, but covers general healthcare rather than AD [12]. The one rigorous systematic critique of QML in digital health concluded that there is no consistent evidence of empirical quantum utility, but deliberately scoped to electronic health records and excluded the neuroimaging-based AD work that constitutes most of the corpus examined here [9]. No review has applied a comparably critical lens to QML for Alzheimer's disease specifically. Table I summarises this landscape.

TABLE I *Comparison of Existing Reviews*

| Review (focus) | AD-spec. | Stance | Gap left |
|---|---|---|---|
| QML medical imaging | No | Descr. | No AD focus |
| QML healthcare | No | Descr. | Catalogue only |
| QC clinical care | No | Mod. crit. | Not AD-specific |
| Sim. QC neuro disease | Partly | Mod. crit. | Disease-broad |
| QDL neuroinformatics | Partly | Mod. crit. | No study critique |
| QML medicine | Partly | Strong crit. | Not AD-specific |
| QML digital health | No | Strong crit. | Excludes AD imaging |
| This review | Yes | Strong crit. | — (fills gap) |

## III. METHODS

### A. Search strategy

Six bibliographic databases were searched: ACM Digital Library, IEEE Xplore, PubMed, ScienceDirect, Scopus, and Springer Nature. The search combined a quantum-computing concept block with an Alzheimer's/dementia concept block, adapted to each database's syntax and restricted to peer-reviewed journal articles and reviews.

### B. Eligibility criteria

A primary study was included only if it: (1) addressed AD or MCI as a separately evaluated diagnostic target; (2) applied a genuine quantum method using qubits, quantum circuits, or quantum kernels; (3) reported at least one quantitative performance metric; (4) was a peer-reviewed journal article; (5) was published 2017–2026; (6) was in English; and (7) had an available full text. Studies were never excluded for being methodologically weak — quality is the object of this review, not an eligibility filter.

### C. Study selection and extraction

The database searches returned 865 records (ACM 303, IEEE Xplore 41, PubMed 21, ScienceDirect 315, Scopus 109, Springer Nature 76). Removing 38 duplicates (33 by DOI, 5 by title) left 827 records. Screening excluded 724 and retained 103: 22 reviews and 81 primary studies. Of the 81, 17 could not be retrieved and 64 were assessed at full text; 49 were excluded (28 wrong disease, 10 wrong type, 7 not genuine quantum computing, 4 no empirical results). Fifteen studies met all criteria. Each was extracted on a 34-field template covering dataset, quantum architecture, execution environment, pre-quantum processing, evaluation protocol, baseline quality, and reported outcome. Fig. 1 presents the PRISMA flow.

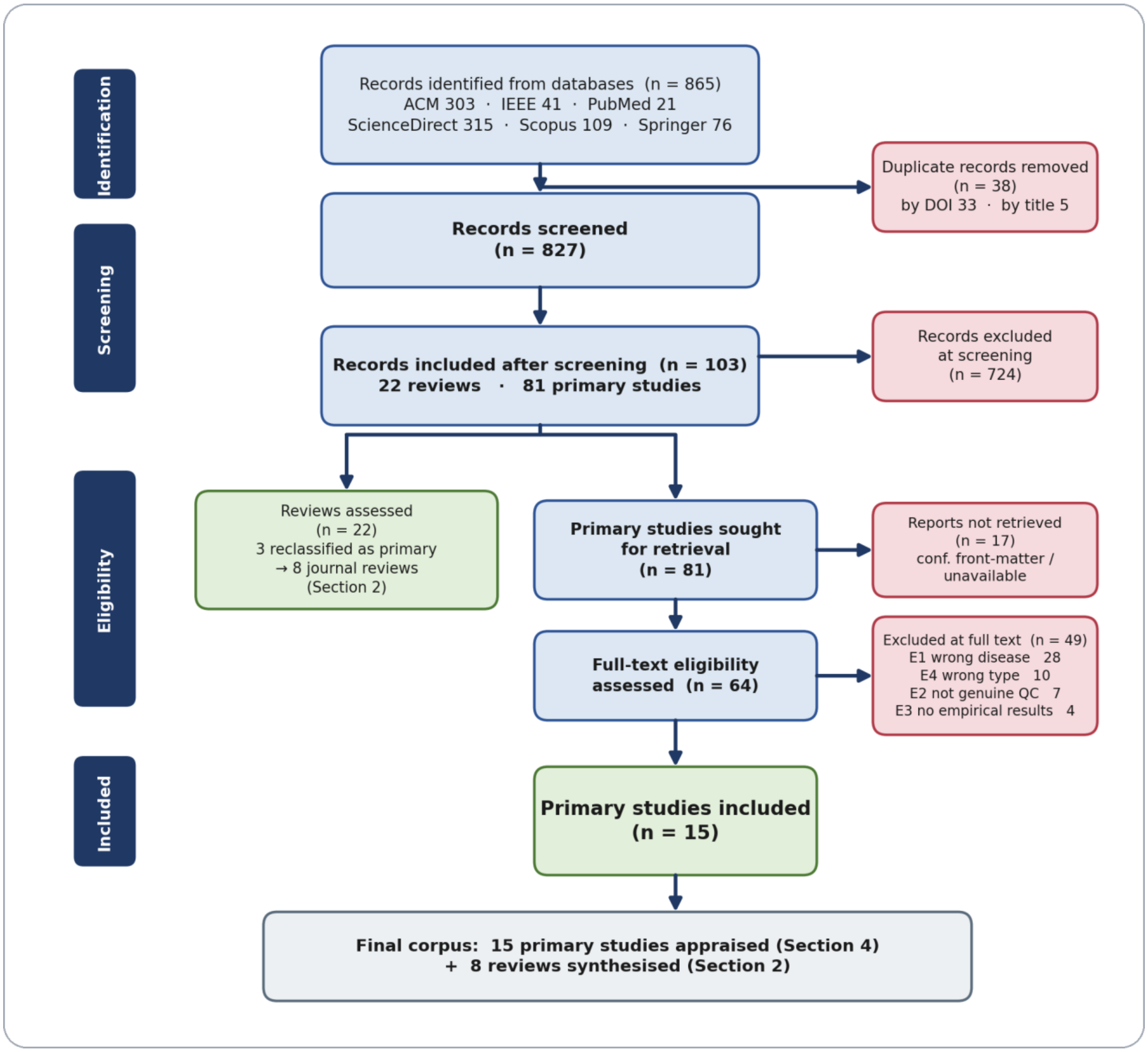


Fig. 1. PRISMA 2020 flow of study selection.

## IV. RESULTS

### *A. Overview of the corpus*

The 15 included studies were published between 2022 and 2026, with 12 in 2025–2026. Eleven used magnetic resonance imaging — ten structural MRI and one resting-state functional MRI — two used EEG, one handwriting dynamics, and one gene-expression data. Quantum architectures varied: hybrid quantum-classical convolutional networks, variational quantum classifiers, quantum support vector machines, and quantum neural networks all appeared. Table II summarises the corpus, and Fig. 2 abstracts the common pipeline.

TABLE II *The 15 Included Primary Studies and Their Principal Characteristics*

| Study (year) | Modality | QML architecture | Dataset / size | Execution | Accuracy |
|---|---|---|---|---|---|
| Amin et al. (2025) | MRI | Hybrid quantum-classical CNN | ADNI-1 / ADNI-1 6400 MRI slice | Simulator | 99.78% |

| Study (year) | Modality | QML architecture | Dataset / size | Execution | Accuracy |
|---|---|---|---|---|---|
| Nagajyothi & Chirra (2025) | MRI | Optimized quantum CNN | OASIS-Kaggle dementia/ Not reported | Simulator | 99.0% |
| Islam et al. (2025) | MRI | Hybrid classical-quantum CNN | OASIS-2 / OASIS-2: 150 subjects | Simulator | 97.5% |
| Belay et al. (2024) | MRI | Ensemble CNN feat. + QSVM | ADNI-1 and ADNI-2 / Not fully reported | Sim. (unclear) | 99.89% |
| Princy et al. (2026) | Genomic | Hybrid CNN + VQC | GEO transcriptomic data / AD: 240 samples | Simulator | 91.0% |
| Ho et al. (2025) | EEG | VQC / QSVM / QNN | Temple University Hosp / 120 subjects | Simulator | 96.58% |
| Radhi et al. (2025) | MRI | Hybrid quantum-classical net | Kaggle Alzheimer's MRI / Kaggle ~5,121 | Simulator | 99.67% |
| Ramesh et al. (2025) | EEG | Graph-attention + hybrid QCNN | CAUEEG / Very small - AD 12, | Simulator | 99.96% |
| Sabari Vasan & Jayalakshmi (2026) | MRI | Quantum deep neural network | The Best Alzheimer's M / 11,519 axial MRI image | Simulator | 98.0% |
| Shankar et al. (2025) | MRI | Classical GNN + QAOA tuning | ADNI / 12,000 MR images | Sim. (unclear) | 98.10% |
| Cappiello & Caruso (2025) | Handwriting | Quantum-kernel SVC (+ QNN) | DARWIN handwriting data / 174 participants | Simulator | 88.29% |
| Kaewta et al. (2026) | MRI | Hybrid quantum CNN (+ QGAN) | AMD dataset / AMD 6400 images | Simulator | 93.0% |
| Choi et al. (2025) | fMRI | Hybrid 1D-CNN + QCNN | ADNI / 365 individuals | Simulator | 58.1%* |
| Shahwar et al. (2022) | MRI | ResNet34 transfer + quantum | Kaggle Alzheimer MRI d / 6400 labelled MRI image | Simulator | 97.2%† |
| Alsharabi et al. (2023) | MRI | AlexNet transfer + quantum | ADNI / AD task | Simulator | 94.0% |

*Accuracy is each study's representative headline figure as a percentage. *Choi et al. report balanced accuracy 0.581 (≈ 58.1%), near chance. †Shahwar et al.'s test accuracy is shown, not the higher training accuracy.*

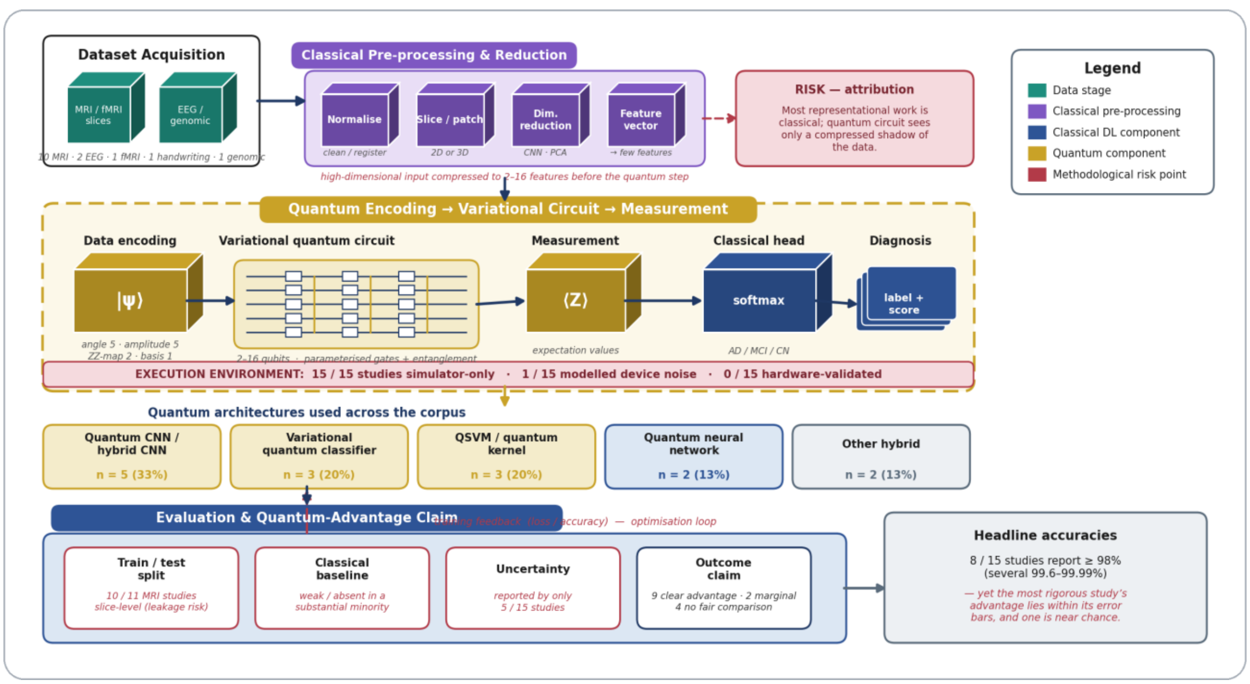


Fig. 2. Generic QML pipeline for AD diagnosis, annotated with corpus statistics. All quantum stages were simulated (15/15), one modelled noise, none were hardware-validated.

*B. Execution environment*

The single most consistent finding concerns where these experiments were run. All 15 studies were executed on classical simulators — most commonly Qiskit Aer or PennyLane default devices. Not one reported unambiguous result on real quantum hardware: 13 explicitly used simulators, one described its backend ambiguously, and one did not specify. Only one study modelled device noise; the remaining 14 assumed noiseless execution or did not address noise. A classifier that performs well on a noiseless simulator has not demonstrated that a physical device would do so, nor that the computation could not be reproduced classically [11]. A literature in which every result is simulator-based cannot by itself establish a quantum advantage.

*C. Datasets, partitioning, and leakage*

Sample sizes were frequently very small — in the most extreme case on the order of thirty subjects across three classes. Fig. 3 shows that near-perfect accuracies are reported across the full range of cohort sizes, and the only near-chance result was obtained under careful subject-level evaluation. Many image-based studies partitioned MRI data at the level of slices rather than subjects; when slices from the same patient appear in both partitions, accuracy is inflated by leakage [16,27], a problem documented specifically for deep-learning studies of AD [4]. Our extraction flagged leakage-prone partitions in the majority of MRI studies.

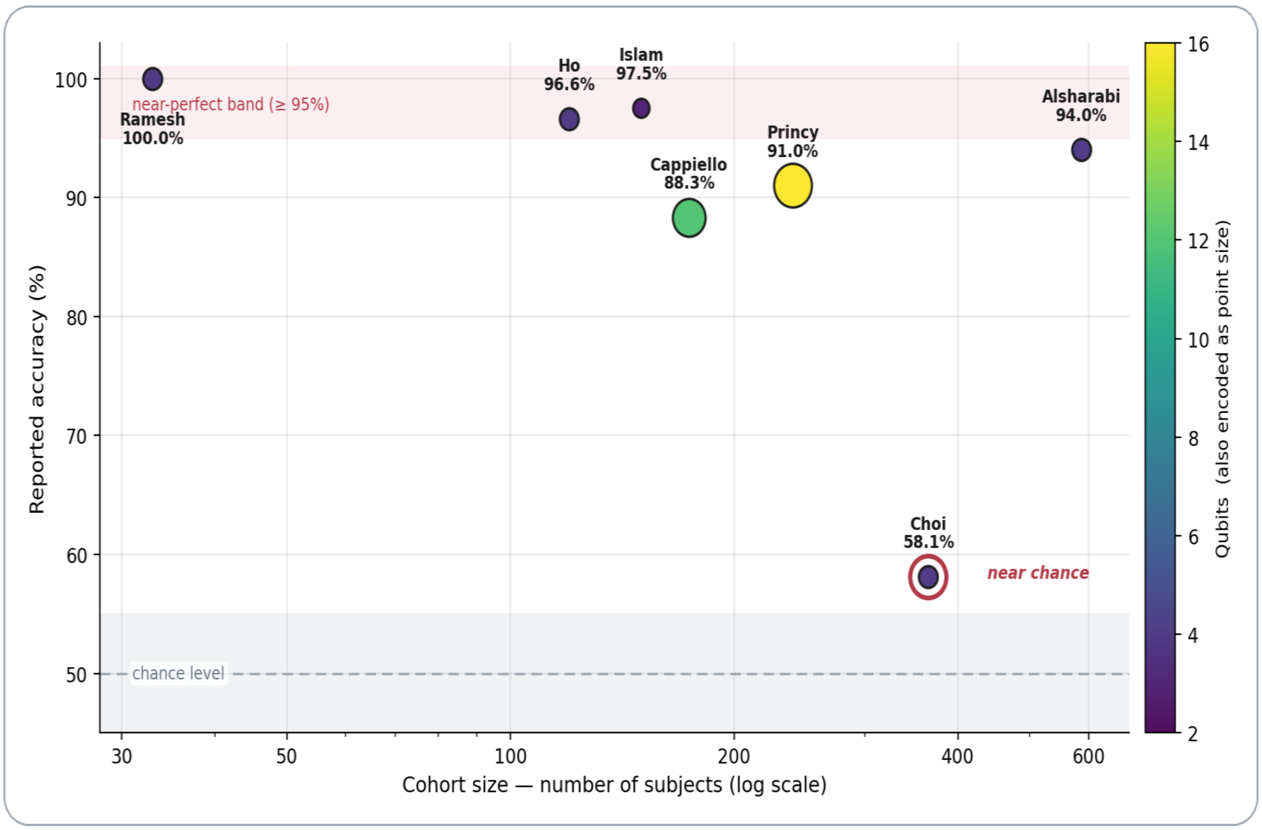


Fig. 3. Reported accuracy vs. cohort size for the seven studies reporting a genuine subject count; point size/colour encode qubit count. The near-chance result (circled) used subject-level evaluation.

*D. Dimensionality reduction*

Reported registers ranged from two to sixteen qubits, with several studies using only two to six. High-dimensional inputs cannot be fed directly into such circuits, so every image- and genomics-based study applied substantial classical pre-processing to compress the input to a few features. In one genomics study, more than twenty thousand features were reduced to sixteen, and in several imaging studies a classical convolutional network reduced images to as few as two to four features — leaving most representational work to classical components.

*E. Baselines, uncertainty, and outcomes*

Baseline quality was uneven. In a substantial minority of studies, the baseline was weak or effectively absent — the only comparison being accuracy figures quoted from other papers. An advantage measured against a weak baseline is not evidence of a quantum advantage [11]. Only five of the 15 studies reported any measure of variability; the remaining ten reported point estimates only. Nine studies claimed a clear advantage and two a marginal one, while four provided no proper classical comparison. Among those reporting a comparison, seven reported accuracies at or above 98%, several reaching 99.6–99.99%.

Two studies illustrate why these numbers should be read cautiously. The most methodologically careful study — using subject-level partitioning and standard deviations across twenty splits — reached 88.3% (±4.7%) against a best classical model at 85.3% (±7.5%); the advantage is smaller than the overlapping standard deviations. Conversely, a study evaluating a hybrid model on a deliberately hard task reported a balanced accuracy of only about 0.58 — barely above chance. The studies that evaluated themselves most rigorously found either a negligible advantage or near-chance performance — the opposite of what a genuine, robust quantum advantage would produce.

*F. Synthesis*

The corpus consists of simulator-only experiments, almost never noise-modelled, frequently trained on small and potentially leakage-prone datasets, heavily dependent on classical pre-processing, and often compared against weak baselines without uncertainty quantification. Fig. 4 summarises this profile; on every dimension except adequate baselines, the corpus is dominated by studies that do not meet the standard a credible claim would require. Fig. 5 locates each risk at the pipeline stage where it is introduced. On the evidence assembled here, the AD-specific QML literature does not currently demonstrate a quantum advantage for Alzheimer's diagnosis [9].

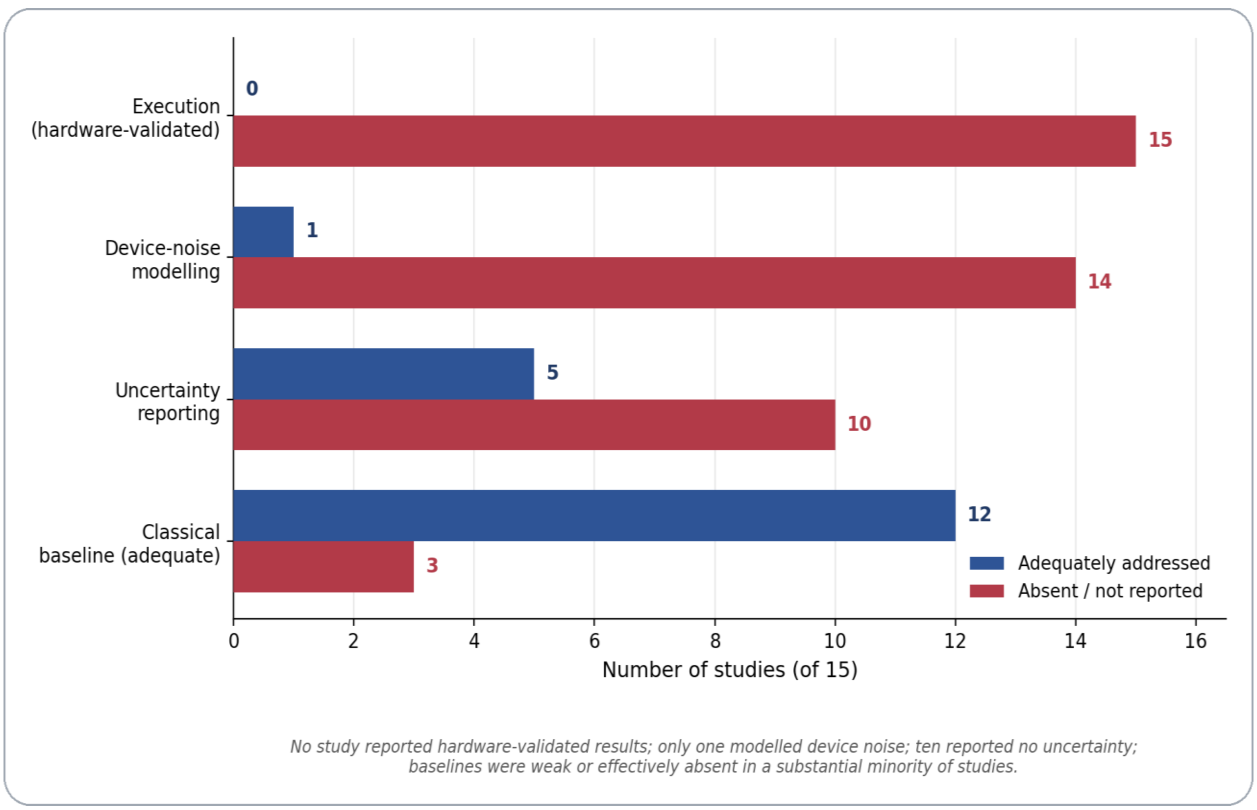


Fig. 4. Methodological profile of the 15-study corpus across the four dimensions most relevant to a quantum-advantage claim.

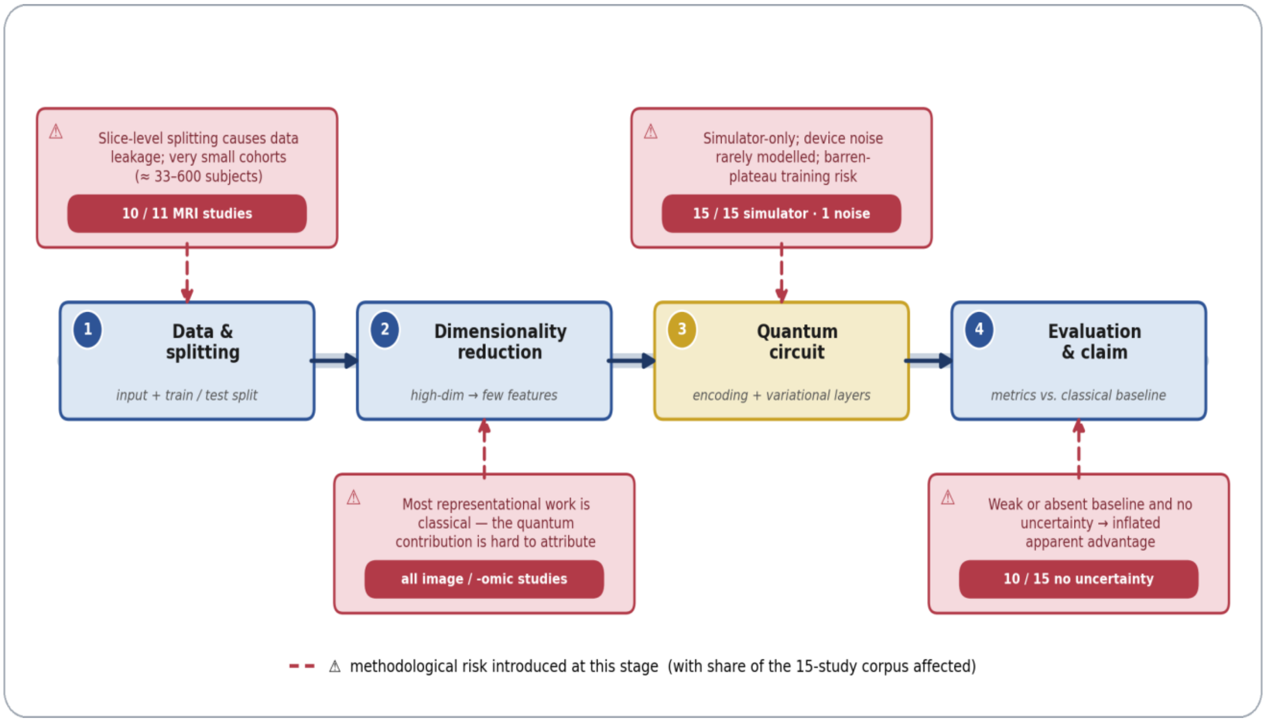


Fig. 5. Where methodological risk enters the QML-for-AD pipeline; each marker indicates the stage at which a threat to validity is introduced.

## V. DISCUSSION

It should be stated clearly what this conclusion is not. It is not a claim that quantum methods cannot help with AD diagnosis, nor a criticism of exploratory work. It is a claim about evidential status: the current literature establishes feasibility and motivates further work, but does not yet license the strong quantum-advantage claims that many individual papers make.

### A. *Why the claims outrun the evidence*

Several structural factors explain the gap. Publication incentives favour positive results; the convenience of slice-distributed MRI datasets makes leakage-prone partitioning easy to overlook [4]; small qubit counts make heavy classical pre-processing unavoidable, blurring the classical/quantum boundary; and accessible noiseless simulators make it natural to report simulator results without flagging their distance from hardware reality. None of these implies bad faith, but together they produce a literature that systematically overstates the maturity of the evidence.

### B. *A reporting standard for QML-for-AD studies*

Future studies should state the execution environment explicitly (simulator or hardware, with noise model or device and qubit count); partition at the subject level; report statistical uncertainty for every headline metric; compare against a fair classical baseline of comparable capability evaluated under an identical pipeline; isolate the quantum contribution with an ablation; and provide a full circuit specification with, where possible, public code and data. Adopting these practices would ensure that a reported advantage, if found, is credible — and that a reported null result is informative.

### C. *Limitations*

The search was restricted to peer-reviewed journal articles in English. The corpus of 15 studies is small, reflecting the youth of the field. Extraction involved reviewer judgement, mitigated by an explicit 34-field template. We deliberately do not pool accuracies into a meta-analytic estimate, which given the heterogeneity and leakage concerns would be misleading rather than informative.

## VI. CONCLUSION

This systematic critical review of 15 peer-reviewed primary studies finds that quantum-advantage claims for AD diagnosis are not, at present, supported by credible evidence. The corpus is simulator-only, almost never models device noise, frequently relies on small and potentially leakage-prone datasets, depends heavily on classical pre-processing, and often compares against weak baselines without quantifying uncertainty. These are largely fixable problems of design and reporting. Until they are addressed, quantum-advantage claims in this field should be regarded as hypotheses awaiting rigorous test, not as established results.

REFERENCES


[1] M. K. Awang, G. Ali, and M. Faheem, “Recent advancements in neuroimaging-based Alzheimer’s disease prediction using deep learning approaches in e-Health: a systematic review,” Health Sci. Rep., vol. 8, no. 5, e70802, 2025.

[2] J. Amin, M. U. Ali, M. Z. Islam, and S. W. Lee, “Quantum AI for psychiatric diagnosis: enhancing dementia classification with quantum machine learning,” Front. Psychiatry, vol. 16, 1648060, 2025.

[3] M. K. Awang et al., “Neuroimaging-based AD prediction using deep learning: a systematic review,” Health Sci. Rep., vol. 8, no. 5, e70802, 2025.

[4] P. Bhattarai et al., “Data leakage in deep learning for Alzheimer’s disease diagnosis: a scoping review,” Diagnostics, vol. 15, no. 18, 2348, 2025.

[5] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, “Quantum machine learning,” Nature, vol. 549, no. 7671, pp. 195–202, 2017.

[6] G. Cappiello and F. Caruso, “Quantum AI for Alzheimer’s disease early screening,” Neurocomputing, vol. 647, 130565, 2025.

[7] J. Choi et al., “Early-stage detection of cognitive impairment by hybrid quantum-classical algorithm using resting-state fMRI time-series,” Knowl.-Based Syst., vol. 310, 112922, 2025.

[8] S. C. Fairburn et al., “Applications of quantum computing in clinical care,” Front. Med., vol. 12, 1573016, 2025.

[9] H. Gupta et al., “A systematic review of quantum machine learning for digital health,” npj Digit. Med., vol. 8, 115, 2025.

[10] T. K. K. Ho et al., “EEG-based dementia classification using CS-EMD synchrony features and quantum machine learning,” IEEE Trans. Consum. Electron., 2025.

[11] H.-Y. Huang et al., “Power of data in quantum machine learning,” Nat. Commun., vol. 12, 2631, 2021.

[12] R. Idzikowski et al., “A survey on quantum machine learning applications in medicine and healthcare,” Appl. Sci., vol. 16, no. 3, 1630, 2026.

[13] M. Islam, M. J. Hasan, and M. R. C. Mahdy, “CQ-CNN: a lightweight hybrid classical–quantum convolutional neural network for Alzheimer’s disease detection using 3D structural brain MRI,” PLOS ONE, vol. 20, no. 9, e0331870, 2025.

[14] K. Jenish, S. Keerthana, and R. Karthik, “A review of simulated quantum computing applications in neurological disease detection and classification,” Arch. Comput. Methods Eng., 2026.

[15] C. Kaewta et al., “QENNA: a quantum-enhanced neural network for early Alzheimer’s detection using MRI,” Artif. Intell. Med., vol. 172, 103322, 2026.

[16] S. Kapoor and A. Narayanan, “Leakage and the reproducibility crisis in machine-learning-based science,” Patterns, vol. 4, no. 9, 100804, 2023.

[17] A. Khan et al., “Applications of deep learning in Alzheimer’s disease: a systematic literature review,” Artif. Intell. Rev., vol. 57, no. 11, 310, 2024.

[18] M. Larocca et al., “Barren plateaus in variational quantum computing,” Nat. Rev. Phys., vol. 7, pp. 174–189, 2025.

[19] Y. Liu, S. Arunachalam, and K. Temme, “A rigorous and robust quantum speed-up in supervised machine learning,” Nat. Phys., vol. 17, no. 9, pp. 1013–1017, 2021.

[20] D. Nagajyothi and V. R. R. Chirra, “Optimized quantum CNN with improvised iHow algorithm for MRI-based dementia diagnosis,” IEEE Access, vol. 13, 2025.

[21] N. A. Orka et al., “Quantum deep learning in neuroinformatics: a systematic review,” Artif. Intell. Rev., vol. 58, 136, 2025.

[22] M. J. Page et al., “The PRISMA 2020 statement: an updated guideline for reporting systematic reviews,” BMJ, vol. 372, n71, 2021.

[23] P. R. Polu, “AI-driven chemometrics for multi-omics data integration: advances, challenges, and future directions,” Crit. Rev. Anal. Chem., 2026.

[24] R. Princy et al., "Hybrid quantum-machine learning models for high-dimensional gene expression profiling and early risk prediction of Parkinson's and Alzheimer's disease," Genet. Mol. Res., vol. 25, no. 1, 2026.

[25] M. M. Radhi et al., "Hybrid quantum-assisted deep learning model for early-stage Alzheimer's disease classification based on MRI images," Mesopotamian J. Big Data, vol. 2025, pp. 156–177, 2025.

[26] S. Ramesh et al., "Enhanced Alzheimer's disease identification from central lobe EEG using multi-aspect quantum-classical graph attention networks," Biomed. Mater. Devices, 2025.

[27] M. Rosenblatt, L. Tejavibulya, R. Jiang, S. Noble, and D. Scheinost, "Data leakage inflates prediction performance in connectome-based machine learning models," Nat. Commun., vol. 15, 1829, 2024.

[28] S. Sabari Vasan and P. Jayalakshmi, "Alzheimer's disease detection using a quantum deep neural network with Haralick feature extraction and simulated annealing optimization," PeerJ Comput. Sci., vol. 12, e3387, 2026.

[29] T. Shahwar et al., "Automated detection of Alzheimer's via hybrid classical quantum neural networks," Electronics, vol. 11, no. 5, 721, 2022.

[30] V. G. Shankar, D. S. Sisodia, and P. Chandrakar, "Alzheimer's stage progression modeling using graph neural network and MRI biomarkers," Neural Comput. Appl., vol. 37, pp. 16825–16847, 2025.

[31] M. Swan, R. P. dos Santos, and F. Witte, "Quantum neurobiology," Quantum Rep., vol. 4, no. 1, pp. 107–126, 2022.

[32] U. Ullah and B. Garcia-Zapirain, "Quantum machine learning revolution in healthcare: a systematic review," IEEE Access, vol. 12, pp. 11423–11450, 2024.

[33] L. Wei, H. Liu, J. Xu, L. Shi, Z. Shan, B. Zhao, and Y. Gao, "Quantum machine learning in medical image analysis: a survey," Neurocomputing, vol. 525, pp. 42–53, 2023.