International Journal of Scientific Modern Research and Technology
(Volume: 24, Issue: 02, Number: 01) Paper ID: IJSMRT - 26080201
ISSN: 2582-8150

Quantum Machine Learning Approaches for Early Detection and Clinical Progression Classification of Progressive Supranuclear Palsy and Parkinson's Disease: A Critical Review

Kumar Sanu[1], Madhuvan Dixit[2]
[1]M. Tech Scholar, Head and Professor
Department of IT, NIIST, Bhopal, India

Abstract: Progressive supranuclear palsy (PSP) and Parkinson's disease (PD) share bradykinesia, rigidity, gait dysfunction, and non-motor manifestations, yet differ substantially in molecular pathology, prognosis, therapeutic response, and trial eligibility. Early PSP-parkinsonism may therefore be labeled as PD until vertical gaze abnormalities, early falls, axial rigidity, or characteristic imaging changes become evident. Classical machine learning can integrate clinical scales, magnetic resonance imaging (MRI), molecular biomarkers, speech, and wearable signals, but performance estimates are often weakened by small samples, spectrum bias, data leakage, and limited external validation. Quantum support vector machines (QSVMs) offer a different kernel-construction mechanism: classical observations are encoded into quantum states and pairwise similarities are estimated in a high-dimensional Hilbert space. This critical review evaluates whether that mechanism is presently capable of improving early differential diagnosis and longitudinal progression classification. A structured search of biomedical and engineering literature was organized around PSP/PD diagnosis, progression biomarkers, machine learning, quantum kernels, and quantum health applications. Evidence supports MRI-derived midbrain and superior cerebellar peduncle measures, MR Parkinsonism Index variants, multimodal clinical data, and digital motor features as candidate inputs. However, no mature, prospectively validated QSVM pipeline for PSP-versus-PD classification was identified. Existing biomedical QML studies are predominantly small, retrospective demonstrations, and quantum advantage over strong classical baselines remains unproven. We propose a leakage-safe hybrid architecture and a staged validation framework emphasizing nested cross-validation, site-held-out testing, calibration, decision-curve analysis, hardware-aware kernel auditing, and prospective silent deployment. QSVM should currently be treated as an experimentally testable representation strategy—not a clinically established diagnostic technology. The near-term scientific opportunity lies in rigorously determining when quantum kernels add information beyond classical kernels in low-sample, multimodal precision-neurology tasks.

Keywords: atypical parkinsonism, biomarkers, clinical progression, magnetic resonance parkinsonism index, Parkinson's disease, progressive supranuclear palsy, quantum kernel, quantum machine learning, quantum support vector machine.

How to cite this article: Kumar Sanu, Madhuvan Dixit. (2026). Quantum Machine Learning Approaches for Early Detection and Clinical Progression Classification of Progressive Supranuclear Palsy and Parkinson's Disease: A Critical Review, International Journal of Scientific Modern Research and Technology (IJSMRT), ISSN: 2582-8150, Volume-24, Issue-02, Number-01, Aug-2026, pp. 01-12, URL: https://www.ijsmrt.com/wp-content/uploads/2026/09/IJSMRT-26080201.pdf

IJSMRT-26080201

I. INTRODUCTION

Parkinsonian syndromes are defined clinically by combinations of bradykinesia, rigidity, rest tremor, and postural instability, but the same surface phenotype may arise from distinct proteinopathies. PD is primarily associated with degeneration of nigrostriatal dopaminergic neurons and α-synuclein pathology, whereas PSP is a four-repeat

tauopathy with neuronal and glial tau deposition involving basal ganglia, diencephalic, brainstem, and cortical networks [1]–[4]. This distinction matters because PSP typically progresses more rapidly, levodopa responsiveness is limited or transient, falls and dysphagia occur earlier, and counseling, supportive interventions, and clinical-trial pathways differ.

The Movement Disorder Society criteria recognize multiple PSP phenotypes and levels of diagnostic certainty, expanding beyond classic Richardson syndrome [1]. This improves sensitivity to early and variant presentations but also exposes an unavoidable longitudinal problem: a patient may initially satisfy a PD-like phenotype and only later manifest vertical supranuclear gaze palsy, slow vertical saccades, early postural instability, frontal-executive dysfunction, or axial-predominant rigidity. Longitudinal work has shown that a subset initially diagnosed with PD later develops features consistent with PSP-parkinsonism, illustrating why a single cross-sectional examination can be insufficient [5].

MRI can reveal midbrain atrophy, third-ventricle enlargement, changes in the pons-to-midbrain relationship, and superior cerebellar peduncle involvement. The magnetic resonance parkinsonism index (MRPI) and MRPI 2.0 formalize several of these measurements, while automated segmentation, diffusion measures, and radiomics broaden the feature space [5]–[10]. Fluid neurofilament light, glial fibrillary acidic protein, tau-related measures, dopaminergic imaging, FDG-PET covariance patterns, speech, eye movements, and wearable-derived gait variables provide complementary information [11]–[17]. The statistical challenge is not a shortage of candidate variables; it is learning reproducible structure from heterogeneous, correlated, partly missing data collected in comparatively small PSP cohorts.

Support vector machines (SVMs) are attractive in such settings because regularization and kernel methods can work well when the number of features is large relative to the number of observations. Deep networks can learn richer representations from raw images or time series, but their data requirements and calibration behavior may be unfavorable in rare-disease cohorts. Quantum kernel methods retain the margin-based SVM framework while replacing a conventional kernel with similarities generated by quantum state overlaps [18]–[24]. The approach is scientifically interesting because quantum feature maps may induce representations that are difficult to reproduce classically. Yet expressivity does not guarantee generalization, and a computationally difficult kernel is not automatically clinically useful.

This review asks three questions. First, which clinical and biomarker signals are sufficiently plausible and standardized to form an early PSP/PD feature set? Second, what does present evidence actually establish about QSVMs in biomedical classification? Third, how should a clinically meaningful experiment be designed so that apparent quantum benefit is not an artifact of leakage, weak baselines, class imbalance, or repeated tuning on the test set? The contribution is deliberately critical: it connects neurology and quantum learning while preserving the distinction between demonstrated performance and a proposed research program.

II. REVIEW METHODOLOGY

A structured critical-review strategy was used rather than a statistical meta-analysis because the target literature is heterogeneous and direct clinical QSVM studies in PSP are sparse. Searches were conceptually organized across PubMed/biomedical indexes, IEEE Xplore, and multidisciplinary databases using combinations of “progressive supranuclear palsy,” “Parkinson disease,” “early diagnosis,” “progression,” “MRI,” “MRPI,” “radiomics,” “wearable,” “machine learning,” “quantum kernel,” “quantum support vector machine,” and “quantum machine learning healthcare.” Priority was given to diagnostic criteria, longitudinal or pathology-supported studies, multicenter imaging work, systematic reviews, foundational quantum-kernel theory, hardware experiments, and comparative biomedical demonstrations published through August 2026.

Records were considered relevant when they addressed: (i) differentiation of PSP, PD, healthy controls, or other parkinsonian syndromes; (ii) progression or severity measurement; (iii) computational biomarkers potentially usable as model inputs; (iv) mathematical or empirical properties of quantum kernels; or (v) biomedical QML

evaluation. Purely promotional perspectives without testable methods were used only for contextual framing. Because the literature does not support a pooled estimate of QSVM accuracy for PSP-versus-PD, no summary sensitivity or specificity was calculated. Instead, evidence was synthesized by modality, validation maturity, and risk of bias. This method avoids a faux meta-analysis whose pooled decimal places would be more precise than the underlying science.

The review has three limitations. Database coverage was structured but not registered prospectively as a PRISMA systematic review; rapidly evolving preprints and post-August 2026 papers may alter the QML landscape; and the absence of reported direct evidence cannot prove that no unpublished study exists. These constraints are reflected in cautious conclusions and in the use of a proposed validation framework rather than performance claims.

III. CLINICAL PROBLEM DEFINITION

A. Overlap and Differential Features

Table 1: Clinical and Modeling Contrasts between PD and PSP [1]–[5]

Dimension	Parkinson's disease	Progressive supranuclear palsy	Modeling implication
Core pathology	α-synucleinopathy; nigrostriatal degeneration	Four-repeat tauopathy with neuronal and glial inclusions	Labels based only on early phenotype can be noisy
Motor pattern	Often asymmetric; rest tremor common; limb-predominant early	Often axial-symmetric; early instability; tremor less typical	Laterality and axial/limb ratios may be informative
Oculomotor signs	Usually preserved early	Slow vertical saccades or vertical gaze palsy	Eye-tracking can detect pre-binary abnormalities
Falls and gait	Often later	Early unexplained falls and freezing phenotypes	Time-to-event modeling may outperform static labels
Levodopa response	Frequently substantial early	Often poor, partial, or transient	Medication-state documentation is essential
Cognition/behavior	Variable; may emerge later	Frontal-executive, apathy, impulsivity may occur early	Cognitive features add signal but raise missingness
MRI	Subtle conventional changes early	Midbrain/SCP and ventricular changes; elevated MRPI	Quantitative automation reduces reader dependence
Progression	Usually slower and heterogeneous	Often faster, phenotype-dependent	Ordinal or survival endpoints need diagnosis-specific design

The diagnostic task should not be reduced to an ahistorical three-class prediction among healthy control, PD, and PSP. Clinically useful targets include: distinguishing early PSP-parkinsonism from PD; estimating the probability of conversion from an initially uncertain parkinsonian presentation; predicting near-term falls, dysphagia, cognitive decline, or loss of independence; and assigning progression strata with calibrated uncertainty. Each target requires a different label, sampling frame, and temporal anchor. Mixing them creates a model that is easy to publish and hard to interpret.

B. Spectrum and Label Bias

Pathologic confirmation is the strongest disease label but is rarely available during model development. Consensus clinical diagnosis at longitudinal follow-up is more feasible, though circularity arises when MRI features

incorporated into clinical adjudication are also used as predictors. Models trained on classic PSP-Richardson syndrome and typical tremor-dominant PD may obtain impressive separation while failing exactly where support is needed: mild, variant, or uncertain disease. A robust study should therefore report performance by PSP phenotype, disease duration, age, sex, medication state, acquisition site, and diagnostic-certainty level, and should include other mimics when intended for real referral populations.

IV. BIOMARKERS AND DATA MODALITIES

A. Structural and Molecular Imaging

The characteristic MRI "hummingbird" or "morning glory" appearance is specific in an appropriate context but insensitive in early or variant disease. Quantitative ratios improve reproducibility. MRPI combines pons and midbrain areas with middle and superior cerebellar peduncle widths; MRPI 2.0 adds ventricular measures intended to improve PSP-parkinsonism discrimination [5]–[8]. Automated deep-learning pipelines can segment relevant structures and reduce manual effort [7], [8]. Radiomics and diffusion features may capture intensity, texture, and tract abnormalities not represented by simple morphometry, with recent multicenter work reporting strong classical discrimination [9], [10]. Such findings are promising, but performance depends on scanner harmonization, segmentation quality, and independent validation.

FDG-PET spatial covariance patterns and tau-targeted PET have also been studied in PSP [12]–[14]. Tau tracer off-target binding and imperfect specificity remain concerns. Dopamine transporter imaging can support degenerative parkinsonism but is less suited to separating PD from PSP because both may show presynaptic dopaminergic deficit. Imaging variables are therefore best treated as complementary rather than individually definitive. For a QSVM, compressed radiomic signatures, region-level standardized uptake values, or learned latent embeddings are more realistic inputs than millions of raw voxels because current qubit counts and sampling costs strongly limit direct amplitude-rich representations.

B. Fluid, Clinical, and Digital Biomarkers

Neurofilament light reflects axonal injury and is often higher in atypical parkinsonism than PD, while GFAP and tau-related measures may add disease or severity information [11]. α-Synuclein seed-amplification assays increasingly reshape biological definition of PD, although assay availability, specimen handling, and phenotype-specific sensitivity must be considered. Clinical scales—including MDS-UPDRS, PSP Rating Scale, cognition, activities of daily living, falls, and medication response—remain essential because a diagnostic algorithm should predict outcomes that matter to patients, not only molecular categories.

Digital biomarkers can convert intermittent clinic visits into dense longitudinal observations. Inertial sensors quantify stride variability, turning, freezing, sway, and movement smoothness; smartphones capture tapping, voice, and activity; eye trackers quantify saccade latency and velocity; and acoustic analysis can characterize dysarthria [15]–[17]. Their strengths are frequency and ecological validity, but device heterogeneity, adherence, contextual confounding, and demographic effects can dominate subtle disease signals. A quantum feature map cannot rescue poor standardization of measurement. Quality control and transparent missing-data handling remain upstream obligations.

Table 2: Multimodal Inputs for Early Differential Diagnosis.

Modality	Candidate variables	Strength	Principal risk	QSVM-ready representation
Structural MRI	MRPI/MRPI 2.0, volumes, cortical thickness, radiomics	Anatomically interpretable	Scanner/site and segmentation effects	10–40 harmonized regional features

 ISSN: 2582-8150

PET/SPECT	Metabolic patterns, tau signal, DAT binding	Molecular/network sensitivity	Cost, tracer specificity, limited samples	Pattern scores or ROI summaries
Fluid/omics	NfL, GFAP, α-synuclein assays, genetics	Biological proximity	Batch effects and missingness	Normalized selected panel
Clinical	MDS-UPDRS, PSPRS, cognition, falls	Direct clinical meaning	Observer and medication effects	Domain scores plus time variables
Wearables	Gait, turning, sway, activity	Dense longitudinal sampling	Device/context drift	Aggregated robust statistics
Speech/eye movement	Prosody, articulation, saccade velocity	Early motor-control sensitivity	Language/hardware confounding	Compact embedding or biomarkers

V. CLASSICAL MACHINE LEARNING BASELINE

A QSVM study is scientifically uninterpretable without strong classical comparators. Minimum baselines include regularized logistic regression, linear and radial-basis-function SVMs, random forest or gradient boosting, and—when sample size permits—a neural architecture appropriate to the modality. Feature preprocessing must be learned only from the training fold: imputation, scaling, ComBat-style harmonization, feature selection, dimensionality reduction, and class balancing all leak information when fitted before resampling. Patient-level separation is mandatory for repeated visits, and site-held-out evaluation is preferable to random splitting for multicenter data.

Accuracy is inadequate for rare PSP classification. Reports should include class-specific sensitivity and specificity, balanced accuracy, macro-F1, area under receiver-operating and precision–recall curves, Matthews correlation coefficient, calibration intercept and slope, Brier score, and confidence intervals obtained with patient-level resampling. For progression, appropriate endpoints include ordinal discrimination, time-dependent AUC, concordance index, integrated Brier score, and clinically defined thresholds. Decision-curve analysis can determine whether a model yields net benefit across referral or treatment thresholds rather than merely improving a statistical score.

Explainability should be matched to the model and question. Coefficients, permutation importance, SHAP-style analyses, counterfactual checks, and region-level saliency may identify drivers, but post-hoc explanations are not proof of causal biomarkers. Stability across folds and sites is more informative than a colorful single-patient heat map. The same standard should apply to quantum models: sensitivity to input perturbation, ablation of feature-map components, kernel-target alignment, and prototype or support-vector analysis can make behavior auditable.

VI. QUANTUM SUPPORT VECTOR MACHINES

A. Mathematical Formulation

For training observations (x_i, y_i), $y_i \in \{-1,+1\}$, a soft-margin SVM solves $\min_{(w,b,\xi)} \tfrac{1}{2}\|w\|^2 + C\Sigma_i\xi_i$ subject to $y_i(w^T\varphi(x_i)+b) \geq 1-\xi_i$ and $\xi_i \geq 0$. The kernel trick replaces explicit inner products by $k(x_i,x_j)=\langle\varphi(x_i),\varphi(x_j)\rangle$. In a fidelity quantum kernel, a parameterized encoding circuit $U\varphi(x)$ prepares $|\varphi(x)\rangle=U\varphi(x)|0\rangle\otimes^n$ and estimates $KQ(x,z)=|\langle\varphi(x)|\varphi(z)\rangle|^2$. The resulting kernel matrix is passed to a conventional SVM optimizer; "quantum" therefore describes representation and similarity estimation, not necessarily the entire training pipeline [18]–[23].

Angle encoding maps scaled features to single-qubit rotations; data re-uploading repeats encoding and trainable transformations; entangling feature maps introduce cross-feature interactions; and projected kernels measure selected observables to reduce concentration. Circuit depth, qubit connectivity, feature ordering, scaling,

entanglement topology, measurement shots, and noise mitigation all become hyperparameters. Because kernel entries require pairwise estimates, naive training costs scale quadratically with sample count in addition to quantum sampling. This is a serious limitation for longitudinal multimodal cohorts and motivates low-rank approximations, landmark sampling, or carefully bounded cohort sizes.

B. Promise and Failure Modes

The theoretical attraction of quantum kernels is access to feature spaces that may be expensive to compute classically. Foundational work demonstrates expressive constructions and identifies tasks with rigorous separation under defined assumptions [18]–[21]. Hardware experiments show that nontrivial high-dimensional kernels can be estimated on noisy processors [22]. Yet data can allow classical learners to approximate targets that are computationally difficult in a direct sense, and the relevant question is generalization advantage on the clinical distribution—not Hilbert-space dimension [20].

Several failure modes are particularly relevant. First, overly expressive feature maps can make kernel matrices close to the identity, memorizing training examples while generalizing poorly. Second, fidelity kernels may exhibit exponential concentration: off-diagonal entries become extremely small and distinguishing them demands many shots [23]. Third, noise can flatten or bias similarities. Fourth, small datasets yield high variance and facilitate inadvertent model selection on test results. Fifth, a quantum model may be compared only with a default RBF-SVM instead of tuned classical kernels, deep embeddings, or projected quantum-kernel surrogates. Finally, simulator results do not establish hardware feasibility or runtime advantage.

Table 3: Fair Comparison of Classical and Quantum-Kernel SVMs

Property	Classical SVM	Quantum-kernel SVM	Required fair comparison
Representation	Explicit features or classical kernel	Quantum circuit-induced state similarity	Match preprocessing and training folds
Hyperparameters	C, γ, kernel family	C plus encoding, depth, scale, entanglement, shots	Nested tuning with equal search budget
Compute	CPU/GPU; mature libraries	Simulator or QPU plus shot/queue cost	Report wall time, calls, shots, energy where possible
Noise	Numerical and sampling variation	Shot noise, gates, readout, drift	Ideal, noisy simulator, and hardware audit
Interpretability	Established kernel diagnostics	Feature-map and kernel diagnostics evolving	Ablation and stability analysis
Clinical maturity	Extensive precedent	Early demonstrations; no established PSP pathway	External and prospective validation

VII. EVIDENCE FOR QML IN BIOMEDICINE AND PARKINSONISM

Biomedical QML reports span cancer, electronic health records, imaging, and tabular risk prediction. Small studies frequently find quantum kernels competitive with classical models, but superiority is inconsistent and often sensitive to feature count, scaling, and split [24]–[28]. A 2025 systematic review of QML in digital health concluded that evidence is promising but methodologically immature, with limited external validation and heterogeneous reporting [24]. Clinical-data experiments on noisy intermediate-scale devices establish feasibility, not clinical utility [25]. Comparative studies often show similar performance between quantum and well-tuned classical kernels [26].

For PD, classical literature is much larger: voice, handwriting, gait, MRI, and clinical data have been classified using SVMs, ensembles, and deep networks [15]–[17], [29]–[31]. Quantum-inspired or quantum-enhanced PD papers are emerging, but direct, independently validated QSVM studies that separate early PSP from PD and

model longitudinal progression were not identified in the evidence set. This is the review's most important negative finding. The title describes a research frontier, not an established clinical technology.

The gap is scientifically useful. PSP cohorts are small enough that margin methods are plausible, while multimodal variables may contain nonlinear interactions that motivate richer kernels. Conversely, the same small-sample setting makes false discoveries easy. A credible study should pre-specify primary endpoints and baselines, lock the external test set, document all exclusions, and publish code, circuit definitions, seeds, and fold assignments. If quantum performance does not exceed classical performance, the result remains valuable when it delineates the regime where additional quantum complexity is unnecessary.

Table 4: Evidence Maturity and Permissible Inference.

Evidence domain	Typical finding	Maturity	Inference for PSP/PD
Foundational quantum kernels	Potential expressive or complexity advantages on selected tasks	Theoretical/engineered	Justifies hypothesis, not clinical benefit
Noisy-hardware kernel experiments	Kernel estimation feasible at small-to-moderate scale	Experimental	Mandates hardware and shot-budget reporting
Biomedical tabular demonstrations	Often competitive; advantage inconsistent	Early retrospective	Use strong baselines and external tests
PD classical ML	Many modalities show discriminative signal	Moderate but heterogeneous	Provides baselines and candidate inputs
PSP imaging ML	MRPI, segmentation, radiomics show promise	Emerging multicenter	Best near-term feature source
Direct QSVM PSP vs PD	No mature prospective validation identified	Evidence gap	Frame as prospective research program

VIII. PROPOSED HYBRID QSVM FRAMEWORK

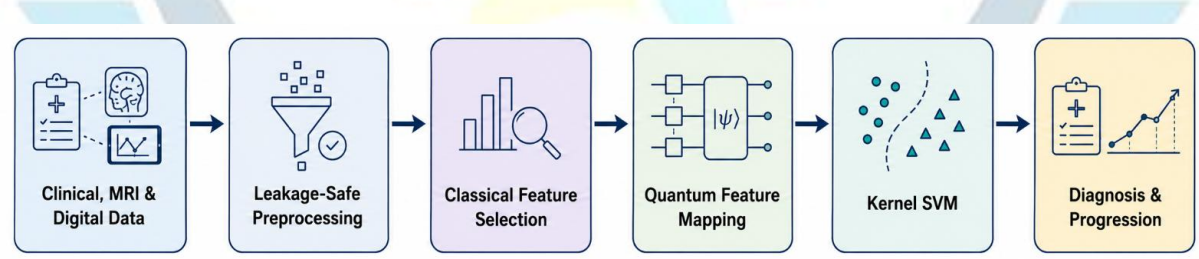

Figure 1: Proposed leakage-safe multimodal quantum-kernel pipeline. Quantum encoding occurs only after fold-specific preprocessing and feature reduction.

A. Cohorts and Endpoints

Development should combine a deeply phenotype PSP cohort with PD and clinically relevant mimic groups. PPMI offers a strong resource for PD progression modeling with longitudinal clinical, imaging, omics, genetic, sensor, and biomarker data, but it does not by itself solve the scarcity of PSP cases [32]. PSP data require multicenter collaboration, harmonized phenotyping, and ideally neuropathologic follow-up. A feasible primary diagnostic endpoint is early PSP/PSP-parkinsonism versus PD at a prespecified disease-duration window, adjudicated after at least two years of follow-up. Secondary endpoints may include PSP phenotype, one-year fall risk, dysphagia progression, and change in PSPRS or MDS-UPDRS domains.

 ISSN: 2582-8150

B. Preprocessing and Feature Engineering

All preprocessing is nested within training data. Continuous variables are robustly scaled; categorical variables are encoded without ordinal assumptions; missingness indicators are considered when clinically meaningful; and multiple imputation is fitted per fold. Site effects in imaging are harmonized using training-site estimates, with sensitivity analysis that preserves biological covariates. Repeated visits from one participant never cross folds. To fit near-term hardware, feature selection should be stability-based and clinically constrained, yielding approximately 8–32 features or a similarly sized latent vector. Feature scaling to circuit rotation ranges must be learned only from training data.

C. Quantum Feature Maps and Kernels

Candidate encodings should include shallow angle-encoding, ZZ-style entangling maps, data re-uploading, and a projected kernel. Circuit families are specified before external testing. Kernel matrices are checked for symmetry, positive semi definiteness, effective dimension, condition number, off-diagonal concentration, and target alignment. Indefinite empirical matrices caused by finite shots may be corrected by a documented eigenvalue procedure, with sensitivity analysis on the uncorrected matrix. Kernel centering and normalization are applied consistently to training and test blocks.

D. Training, Calibration, and Explainability

Nested cross-validation tunes C and circuit parameters. The outer loop estimates internal generalization; a completely held-out site or consortium estimates transportability. Class weights or balanced sampling are selected within the inner loop. Probabilities are calibrated using training-only Platt scaling or isotonic regression. Pairwise multiclass schemes are permissible, but clinically ordered tasks may benefit from hierarchical prediction: degenerative parkinsonism first, PD versus atypical second, PSP versus other atypical syndromes third. Abstention thresholds should be evaluated because uncertain referral is safer than forced confident classification.

Interpretation should include permutation of clinical domains, leave-one-modality-out analysis, circuit-layer ablation, feature-order randomization, and comparison between the quantum kernel and its classical approximations. Patient-level explanations should report uncertainty and plausible measurement changes rather than assert causal pathology. Subgroup performance and calibration across sex, age, disease duration, language, ethnicity, and site are required to detect inequitable error concentration.

IX. VALIDATION AND REPORTING STANDARD

Table 5: Staged Evidence Standard for Translation.

Stage	Question	Minimum evidence	Go/no-go criterion
1. Analytic validity	Is the feature and kernel pipeline reproducible?	Locked code; repeated seeds; simulator checks	Stable kernel and metrics under perturbation
2. Internal validity	Does it generalize within development data?	Nested patient-grouped cross-validation	Benefit beyond uncertainty and tuned baselines
3. External validity	Does it transport across sites/scanners?	Site-held-out independent cohort	Preserved calibration and clinically relevant sensitivity
4. Hardware validity	Does QPU estimation reproduce ideal ranking?	Shot/noise sweep; repeated hardware runs	Acceptable kernel distortion and runtime
5. Clinical utility	Would predictions improve decisions?	Decision curves; workflow simulation	Positive net benefit at prespecified thresholds
6. Prospective validity	Does it work without changing care?	Silent prospective deployment	Stable performance, latency, and subgroup safety

 ISSN: 2582-8150

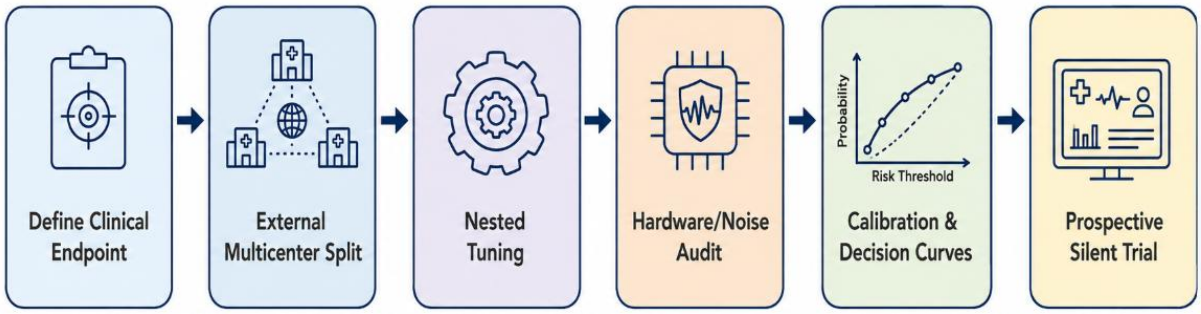

Figure 2: Validation sequence from endpoint specification to prospective silent deployment.

Reporting should follow TRIPOD-AI and related medical-AI principles where applicable, supplemented by quantum-specific details: device and topology; SDK and compiler versions; circuit diagram and gate counts; shots per kernel entry; transpiration strategy; error mitigation; queue and wall-clock time; random seeds; kernel correction; hyperparameter search; and simulator-versus-hardware discrepancies. Dataset statements should describe eligibility, missingness, label adjudication, prevalence, and access restrictions. A model card should state intended users, unacceptable uses, subgroup limitations, and retraining triggers.

Sample-size planning cannot rely on the number of features alone. Precision of sensitivity in the minority PSP class, expected site heterogeneity, number of tuning decisions, and desired external confidence intervals should govern enrollment. Learning curves can determine whether performance is saturation-limited. A priori superiority or noninferiority margins against a clinical-plus-MRPI baseline are preferable to declaring success whenever the QSVM point estimate is numerically larger.

X. ETHICAL, REGULATORY, AND IMPLEMENTATION CONSIDERATIONS

Neurologic diagnoses affect prognosis, employment, driving, caregiving, and psychological well-being. False PSP labeling may cause premature pessimism; false PD labeling may delay anticipatory management of falls, dysphagia, and communication difficulties. The model must therefore be positioned as decision support, with clinician review and explicit uncertainty. Consent should address secondary use of imaging, genomic, speech, and sensor data. De-identification is challenging for MRI and voice, and cross-border federated studies require governance beyond technical privacy claims.

Quantum processing may involve remote cloud infrastructure. Data sent to a quantum service should be minimized and encoded after local preprocessing; contracts should specify retention, jurisdiction, access logs, and incident response. Federated or split workflows may keep raw data at hospitals while sharing model updates or kernel summaries, but privacy leakage remains possible. The novelty of quantum computation does not create an exemption from medical-device regulation, clinical quality management, cybersecurity, or post-deployment monitoring.

Environmental and economic costs also deserve measurement. A small quantum circuit may rely on cryogenic hardware and extensive classical control; simulator-heavy development may consume substantial conventional compute. Comparative studies should report total resource use and latency. A slower, costlier model with the same calibrated clinical utility as an RBF-SVM is not improved merely because its block diagram contains qubits.

XI. OPEN RESEARCH QUESTIONS

1. Which feature distributions produce quantum-kernel geometry that is both non-concentrated and more label-aligned than optimized classical kernels?

2. Can longitudinal change vectors or irregular-time embeddings be encoded efficiently without discarding clinically meaningful temporal structure?
3. Do quantum kernels add value primarily after classical representation learning, or can interpretable clinical variables support an advantage directly?
4. How robust are kernel entries to scanner harmonization, imputation uncertainty, finite shots, device drift, and circuit compilation?
5. Can multicenter PSP consortia establish locked, clinically representative external benchmarks with phenotype and pathology metadata?
6. What level of incremental decision-curve net benefit would justify quantum infrastructure in movement-disorder clinics?

A particularly informative benchmark would release predefined development and site-held-out splits, a clinical-only baseline, MRPI/MRI baselines, multimodal classical kernels, standardized quantum feature maps, and blinded evaluation. Performance would be reported with calibration, uncertainty, runtime, shots, and subgroup analyses. Such a benchmark would convert a fashionable question—"Can quantum AI diagnose PSP?"—into a falsifiable one: "Under which data, circuit, and validation conditions does a quantum kernel improve clinically relevant generalization?"

XII. CONCLUSION

Early differentiation of PSP from PD remains a high-value and technically difficult problem because phenotypes overlap, diagnostic labels evolve, and PSP data are scarce. Quantitative MRI, MRPI variants, radiomics, fluid markers, eye movements, speech, gait, and longitudinal clinical scales provide plausible complementary signals. Classical machine learning already offers strong tools for integrating these data, provided evaluation is leakage-safe, externally validated, calibrated, and clinically anchored.

QSVMs contribute a principled mechanism for constructing nonlinear similarities through quantum feature maps. Foundational and biomedical studies justify careful experimentation, but present evidence does not establish quantum advantage or clinical readiness for PSP/PD classification. No mature prospective QSVM validation at this exact intersection was identified. The responsible conclusion is neither dismissal nor hype: quantum kernels are candidate representations whose incremental value must be tested against optimized classical baselines under realistic site shift, imbalance, noise, and decision thresholds.

The proposed framework offers that test. It prioritizes longitudinally adjudicated labels, compact multimodal representations, nested and site-held-out validation, kernel diagnostics, hardware-aware auditing, calibration, decision curves, fairness analysis, and prospective silent deployment. If those steps demonstrate reproducible benefit, QSVMs may become useful components of precision-neurology pipelines. If they do not, the field will still gain a clearer map of where classical methods remain sufficient—a scientifically successful result in its own right.

REFERENCES

[1] G. U. Höglinger et al., "Clinical diagnosis of progressive supranuclear palsy: The Movement Disorder Society criteria," Movement Disorders, vol. 32, no. 6, pp. 853–864, 2017.

[2] A. L. Boxer et al., "Advances in progressive supranuclear palsy: New diagnostic criteria, biomarkers, and therapeutic approaches," Lancet Neurology, vol. 16, no. 7, pp. 552–563, 2017.

[3] M. Stamelou et al., "Evolving concepts in progressive supranuclear palsy and other 4-repeat tauopathies," Nature Reviews Neurology, vol. 17, pp. 601–620, 2021.

[4] G. G. Kovacs et al., "Distribution patterns of tau pathology in progressive supranuclear palsy," Acta Neuropathologica, vol. 140, pp. 99–119, 2020.

[5] A. Quattrone et al., "Refining initial diagnosis of Parkinson's disease after follow-up: The role of MRPI 2.0," Movement Disorders, vol. 34, 2019.

[6] S. Nigro et al., "Deep learning-based approach for brainstem and ventricular measurement in progressive supranuclear palsy," 2024.

[7] F. Sun et al., "Accurate measurement of magnetic resonance parkinsonism index using a fully automated deep-learning pipeline," 2023.

[8] X. Hu et al., "A deep learning approach to distinguish Parkinson's disease from progressive supranuclear palsy," 2025.

[9] H. Volkmann et al., "MRI classification of progressive supranuclear palsy using artificial-intelligence analysis of structural and microstructural change," 2025.

[10] "A brainstem radiomics framework to distinguish progressive supranuclear palsy from Parkinson's disease," Movement Disorders, 2026.

[11] "Plasma biomarkers and positron emission tomography tau imaging in progressive supranuclear palsy," Movement Disorders, 2023.

[12] R. Buchert et al., "Automatic covariance pattern analysis outperforms visual reading of FDG-PET in variant progressive supranuclear palsy," Movement Disorders, vol. 38, no. 10, pp. 1901–1913, 2023.

[13] G. Marti-Andres et al., "Multicenter validation of metabolic abnormalities related to PSP according to the MDS-PSP criteria," Movement Disorders, vol. 35, pp. 2009–2018, 2020.

[14] J. Jin et al., "Tau PET imaging in progressive supranuclear palsy: A systematic review and meta-analysis," Journal of Neurology, vol. 270, pp. 2451–2467, 2023.

[15] C. Palmisano et al., "Gait initiation in progressive supranuclear palsy: Brain metabolic correlates," Neuro Image: Clinical, vol. 28, art. 102408, 2020.

[16] A. Mirelman et al., "Gait impairments in Parkinson's disease," Lancet Neurology, vol. 18, pp. 697–708, 2019.

[17] C. R. Artusi et al., "Digital progression biomarkers in Parkinson's disease," Movement Disorders, 2021.

[18] V. Havlíček et al., "Supervised learning with quantum-enhanced feature spaces," Nature, vol. 567, pp. 209–212, 2019.

[19] M. Schuld and N. Killoran, "Quantum machine learning in feature Hilbert spaces," Physical Review Letters, vol. 122, art. 040504, 2019.

[20] H.-Y. Huang et al., "Power of data in quantum machine learning," Nature Communications, vol. 12, art. 2631, 2021.

[21] Y. Liu, S. Arunachalam, and K. Temme, "A rigorous and robust quantum speed-up in supervised machine learning," Nature Physics, vol. 17, pp. 1013–1017, 2021.

[22] E. Peters et al., "Machine learning of high dimensional data on a noisy quantum processor," npj Quantum Information, vol. 7, art. 161, 2021.

[23] S. Thanasilp et al., "Exponential concentration in quantum kernel methods," Nature Communications, vol. 15, 2024.

[24] R. S. Gupta et al., "A systematic review of quantum machine learning for digital health," npj Digital Medicine, 2025.

[25] S. Moradi et al., "Clinical data classification with noisy intermediate-scale quantum computers," Scientific Reports, vol. 12, 2022.

[26] X. Vasques et al., "Application of quantum machine learning using quantum kernel algorithms on multiclass biomedical data," Scientific Reports, vol. 13, 2023.

[27] R. V. Aswiga, S. Sridevi, and B. Indira, "Leveraging quantum kernel support vector machine for breast cancer diagnosis from digital breast tomosynthesis images," Quantum Machine Intelligence, vol. 6, art. 40, 2024.

[28] S. K. Sheoran et al., "Robust evaluation of classical and quantum machine-learning methods," Scientific Reports, 2025.

[29] P. Wu et al., "An effective machine learning approach for identifying Parkinson's disease," IEEE Access, vol. 9, 2021.

[30] I. M. El-Hasnony et al., "An optimized ANFIS model using hybrid metaheuristic algorithms for Parkinson's disease prediction," IEEE Access, vol. 8, 2020.

[31] M. Matsumoto et al., "Differentiating Parkinson's disease and progressive supranuclear palsy using kinematic feature extraction and selection," IEEE Access, 2025.

[32] Parkinson's Progression Markers Initiative, "Study design and access to multimodal longitudinal data," PPMI, accessed Aug. 2026.

[33] J. Biamonte et al., "Quantum machine learning," Nature, vol. 549, pp. 195–202, 2017.

[34] M. Cerezo et al., "Variational quantum algorithms," Nature Reviews Physics, vol. 3, pp. 625–644, 2021.

[35] J. R. McClean et al., "Barren plateaus in quantum neural network training landscapes," Nature Communications, vol. 9, art. 4812, 2018.

[36] J. Kübler, S. Buchholz, and B. Schölkopf, "The inductive bias of quantum kernels," Advances in Neural Information Processing Systems, vol. 34, pp. 12661–12673, 2021.

[37] J. Jäger and R. V. Krems, "Universal expressiveness of variational quantum classifiers and quantum kernels for support vector machines," Nature Communications, vol. 14, art. 576, 2023.

[38] S. Jerbi et al., "Quantum machine learning beyond kernel methods," Nature Communications, vol. 14, art. 517, 2023.

[39] G. Agliardi et al., "Mitigating exponential concentration in covariant quantum kernels," npj Quantum Information, 2025.

[40] A. Ali et al., "Explainable machine learning for medical diagnosis: Principles and practice," IEEE Access, 2023.