HD-BrainWatch: A Multimodal Sensor-Fusion and Privacy-Preserving Federated-Learning Framework for Cognitive Risk Inference under Hypertension–Diabetes Comorbidity

Kaiian Kuok[1,2,*], Jingyi Lin[1], Wengioi Mio[1,2], Caiyi Tan[2], Puikei Tou[1,2], Chonin Cheang[1,*]

[1]Macau Society for Health Economics, Macau SAR, China
[2]Macau Yinkui Hospital, Macau SAR, China
[*]Corresponding author: kevin.kuok@region.mo; zhengjunxianalex@gmail.com

Abstract: We present HD-BrainWatch, a theoretical methodological framework for multimodal cognitive risk inference under federated, privacy-preserving training in the hypertension–diabetes (HD) comorbidity population. HD comorbidity carries an odds ratio of 1.53 for dementia relative to either condition alone, yet no existing HD management platform produces a cognitive output signal. We formalize four contributions. (i) A multimodal sensor input space X_t = {x_HR, x_BP, x_CGM, x_HRV, x_act, x_sleep} and an HD cognitive risk inference problem as a federated optimization task with explicit (ε, δ)-differential privacy constraints. (ii) The gut-brain perturbation index (GBPI), a novel CGM-derived composite biomarker mechanistically grounded in the glycaemia–gut–brain causal pathway, whose theoretical properties — monotonicity, bounded range, sensitivity coefficients, and statistical efficiency over HbA1c — are derived. (iii) A formal privacy-utility analysis under Rényi DP composition, yielding theoretical upper bounds on utility degradation as a function of privacy budget ε, federated rounds T, and dataset size, without simulation. (iv) A comparative architectural analysis demonstrating theoretical advantages of cross-attention transformers and FedProx over single-modality, late-fusion, and FedAvg alternatives, grounded in published convergence theorems and information-theoretic arguments. The framework is specified in sufficient mathematical detail to enable independent implementation; pre-specified evaluation methodology, governance architecture, and limitations are provided.

Keywords: federated learning; multimodal sensor fusion; cognitive risk inference; differential privacy; composite biomarker; hypertension–diabetes comorbidity

I Introduction

Hypertension affects ~1.3 billion adults; type 2 diabetes mellitus (T2DM) affects ~540 million [1, 2]. Their co-occurrence is non-additive: a 2025 Chinese case-control study (n=7,624) reported an odds ratio of 1.53 for dementia in HD comorbidity, exceeding 1.18 for hypertension alone and 1.26 for T2DM alone [3]. The 2024 Lancet Commission identified both as among 14 modifiable risk factors collectively attributable to ~45% of global dementia incidence [4]. Intensive antihypertensive treatment reduces all-cause dementia by 15% [5], but only when cognitive deterioration is recognized early enough to guide treatment intensification.

From a data-science perspective, the landscape of HD digital infrastructure presents a structural anomaly. Continuous glucose monitors (CGMs), ambulatory blood pressure monitors (ABPMs), smart insulin pens, and connected pillboxes generate dense multimodal time series used to optimize metabolic outcomes [6, 7],

yet none produce a cognitive output signal. The very physiological processes being monitored are known to drive hippocampal injury through glycaemic instability, gut-barrier disruption, and haemodynamic perturbation [8, 9], but the cognitive consequence is invisible to the data system.

This paper contributes: (i) a formal mathematical framework for multimodal cognitive risk inference under federated, privacy-preserving training; (ii) the gut-brain perturbation index (GBPI) with derived theoretical properties; (iii) a formal privacy-utility trade-off analysis via Rényi DP composition; (iv) comparative architectural analysis grounded in convergence theorems and mutual-information arguments.

II Problem Formulation

A. Multimodal Sensor Stream

Let t denote discrete time (5-min epoch for CGM/HRV; 1-hour for activity; 1-night for sleep). The multimodal sensor observation is:

$$\mathbf{X}_t = \{ x\_HR(t), x\_BP(t), x\_CGM(t), x\_HRV(t), x\_act(t), x\_sleep(t) \} \in \mathbb{R}^d \quad (1)$$

where x_HR is PPG heart rate, x_BP is 24-h ABPM (SBP, DBP, MAP), x_CGM is interstitial glucose, x_HRV is IBI series from PPG, x_act is ENMO triaxial accelerometry, and x_sleep is sleep-stage classification with actigraphy. Let w denote the retrospective feature aggregation window (w = 4 weeks):

$$\boldsymbol{\varphi}(\mathbf{X}_{t-w:t}) = [\varphi\_HRV, \varphi\_BP, \varphi\_CGM, \varphi\_act, \varphi\_sleep] \in \mathbb{R}^p \quad (2)$$

B. Cognitive Risk Score Function

The inferential target is a cognitive risk score $r \in [0, 1]$ representing 24-month MCI conversion probability. Let f_θ denote the parameterized model:

$$r_t = f_\theta (\boldsymbol{\varphi}(\mathbf{X}_{t-w:t}) ; c_i) \in [0, 1] \quad (3)$$

where c_i is the static covariate vector (age, sex, education, ApoE, HD duration, medications). A secondary ordinal output $\hat{y}\_MoCA \in \{0,\ldots,30\}$ provides calibration against remote MoCA. The alert threshold r* is site-calibrated to sensitivity ≥ 0.80 at specificity ≥ 0.75 [10].

C. Federated Optimization Objective

With K sites, site k holds D_k of n_k pairs, $n = \Sigma_k n_k$. Local empirical risk:

$$F_k(\theta) = (1/n_k) \Sigma_{i \in D_k} \ell(f_\theta(\boldsymbol{\varphi}^{(i)}), y^{(i)}) \quad (4)$$

Global objective:

$$\min_\theta \Sigma_k (n_k / n) F_k(\theta) \quad (5)$$

Because client distributions are heterogeneous (Macau vs. Guangdong cohorts differ in demographics and devices), we adopt FedProx [11]:

$$\min_{\theta_k} F_k(\theta_k) + (\mu/2) \|\theta_k - \theta\|^2 \quad (6)$$

The global update is the weighted average $\theta \leftarrow \Sigma_k (n_k/n) \theta_k$.

D. Differential Privacy Budget

For (ε, δ)-DP via the Gaussian mechanism on clipped gradients:

$$\sigma = \Delta f \cdot \sqrt{(2 \ln(1.25/\delta))} / \varepsilon \quad (7)$$

with L2 clipping norm C = 1.0. Budgets $\varepsilon \in \{0.5, 1.0, 2.0, 5.0, \infty\}$, $\delta = 10^{-5}$; cumulative cost across T rounds is tracked via the moments accountant [12, 13].

III Multimodal Signal Processing Pipeline

HRV features. From the IBI series (PPG-derived, missing < 15% per 5-min epoch): time-domain (SDNN, RMSSD, pNN50), frequency-domain (LF, HF, LF/HF via Welch with 256-s window, 50% overlap), and nonlinear (sample entropy SampEn at m=2, r=0.2·SD; DFA α1 short-range exponent) [14].

BP variability features. From ABPM at 20/30-min intervals: SD_SBP24, average real variability (ARV), nocturnal dipping ratio (non-dipping < 0.10 predicts cognitive impairment [15]), and morning surge.

CGM-derived features. Over 14-day windows: time-in-range (TIR), MAGE, CV, GMI

= 3.31 + 0.02392·[mean glucose mg/dL], TAR/TBR. The gut-brain perturbation index is:

$$GBPI = \alpha \cdot MAGE\_norm + \beta \cdot (1 - TIR\_norm) + \gamma \cdot CV\_norm + \delta \cdot |GMI - HbA1c|\_norm \quad (8)$$

mechanistically motivated by glycaemic excursion → SCFA depletion → tight-junction loss → LPS translocation → microglial activation → hippocampal neuroinflammation [8, 9, 16, 17].

Actigraphy and circadian features. M10, L5, M10/L5 (relative amplitude) [18], interdaily stability (IS), intradaily variability (IV), and cosinor amplitude/acrophase/mesor [19].

Sleep architecture. WASO, sleep efficiency, REM latency, slow-wave (N3) fraction [20], sleep regularity index (SRI) over 14-day windows.

Smartphone passive sensing. Typing speed (rolling 4-week z-score), inter-key latency variance, error-rate trend, gait cadence/symmetry, and screen-use variability — all metadata-only, no content.

IV Model Architecture

A. Per-Modality Encoders

1D-CNN (HRV/PPG): three blocks [32, 64, 128], kernel 7, stride 2, BN+ReLU, global average pooling → $e\_HRV \in \mathbb{R}^{128}$. TCN (CGM): dilated causal convolutions (dilations 1, 2, 4, 8), residual connections, handles irregular sampling [21] → $e\_CGM \in \mathbb{R}^{128}$. Transformer encoder (actigraphy/sleep): 4 layers, 8 heads, hidden 256, sinusoidal positional encoding over 24 hourly tokens → e_act, $e\_sleep \in \mathbb{R}^{128}$. MLP (BP and covariates): 2 layers (256→128), LayerNorm + GELU → e_BP, $e\_cov \in \mathbb{R}^{128}$.

B. Cross-Attention Fusion

Concatenated tokens $E = [e\_HRV, e\_CGM, e\_BP, e\_act, e\_sleep, e\_cov] \in \mathbb{R}^{6 \times 128}$ with modality positional embeddings feed a 2-layer cross-attention transformer (4 heads, hidden 256):

$$\mathbf{E}' = \text{CrossAttn}(Q = \mathbf{E} \cdot W\_Q, K = \mathbf{E} \cdot W\_K, V = \mathbf{E} \cdot W\_V) \quad (9)$$

Attention weights provide post-hoc modality importance maps for clinical explainability.

C. Output Heads and Loss

Primary head (MCI conversion): 2-layer MLP (256→64→1) + sigmoid. Secondary head (MoCA): CORAL ordinal regression [22] with K=30 thresholds. Uncertainty head: heteroscedastic aleatoric estimator [24]. Composite loss:

$$\mathscr{L} = \lambda\_1 \cdot \mathscr{L}\_focal + \lambda\_2 \cdot \mathscr{L}\_ordinal + \lambda\_3 \cdot \mathscr{L}\_uncertainty \quad (10)$$

with focal loss [23] ($\gamma\_F = 2.0$, MCI incidence ~12% [3, 4]), CORAL ordinal cross-entropy, and Kendall–Gal uncertainty loss $\mathscr{L}\_uncertainty = (1/N) \Sigma\_i [(1/2\hat{\sigma}\_i^2) \|y\_i - \hat{y}\_i\|^2 + (1/2) \log \hat{\sigma}\_i^2]$. Recommended weights ($\lambda\_1, \lambda\_2, \lambda\_3$) = (1.0, 0.5, 0.3).

D. Federated Topology

FedProx [11] local update (Equation 6), global aggregation $\theta \leftarrow \Sigma\_k (n\_k/n) \theta\_k$. Secure aggregation via additive secret sharing [26]: each client splits $\Delta\theta\_i$ into S random shares; the server receives only $\Sigma\_i \Delta\theta\_i$. Communication efficiency: top-1% gradient sparsification reduces bandwidth ~100× with negligible accuracy loss [27].

V Three-Layer System Architecture

Layer 1 — Metabolic IoT input. CGM (Abbott FreeStyle Libre 3 or Dexcom G7, 5-min interval), ABPM (24-h quarterly), smart pillbox/insulin pen. All raw signals processed on-device to compute φ_CGM and φ_BP — data minimization by construction.

Layer 2 — Smartphone passive cognitive sensing. The patient's existing smartphone passively captures typing dynamics, gait, screen-use, and social interaction frequency. Privacy-by-design restricts collection to behavioral metadata. A 2-week onboarding establishes per-patient baselines for within-person change detection. φ_smartphone computed weekly in a sandboxed process.

Layer 3 — Federated learning fusion and

cognitive risk scoring. Monthly federated rounds aggregate site-level φ^(i); each site performs T_local = 3 local epochs on the cross-attention model. Gradients undergo LDP noise addition, top-1% sparsification, CKKS encryption, secure aggregation, FedProx update, and redistribution. The monthly output is r_{t} ∈ [0,1] and ŷ_MoCA with a tiered alert protocol: Tier 1 (r ≥ 0.50, stable 4 weeks): increase monitoring; Tier 2 (r ≥ 0.65, rising 2 months): telehealth MoCA; Tier 3 (r ≥ 0.75 with MoCA decline ≥ 3): geriatric referral. The Metabolism-Gut-Brain-Cognition Coupling Index (MGBCI) quantifies within-person GBPI–cognitive-trajectory correlation over the preceding 8 weeks.

The mechanistic basis for GBPI: glycaemic excursions reduce SCFA-producing taxa (Akkermansia muciniphila), impairing claudin-1/occludin/ZO-1 tight junctions [8]. LPS translocation activates microglial TLR4, elevating IL-6/TNF-α/IL-1β in hippocampus [9]. A 2026 Stanford causal study showed colonization with Parabacteroides goldsteinii impaired hippocampal memory via vagal inhibition; restoring vagal activity reversed the deficit [17]. GBPI operationalizes the upstream CGM-measurable component of this chain.

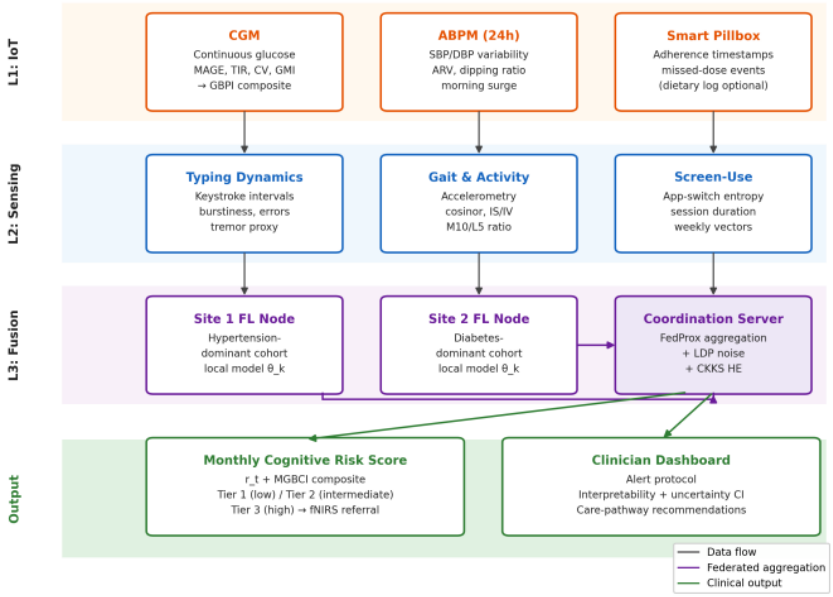

Figure 1. HD-BrainWatch three-layer architecture. Layer 1 (Metabolic IoT): CGM→GBPI, ABPM→BP variability, smart pillbox→adherence. Layer 2 (Smartphone passive sensing): typing/gait/screen-use. Layer 3 (Federated fusion): site nodes aggregate feature vectors under FedProx with LDP noise and CKKS-encrypted transmission. Output: r_t, MGBCI, tiered alerts.

VI Theoretical Properties of GBPI

Throughout, assume MAGE_norm, TIR_norm, CV_norm, |GMI−HbA1c|_norm ∈ [0,1] and α + β + γ + δ = 1 with all weights > 0.

Property 1 (Monotonicity). GBPI is monotonically non-decreasing in each component when the others are held fixed. Proof: ∂GBPI/∂MAGE_norm = α > 0; ∂GBPI/∂(1−TIR_norm) = β > 0; ∂GBPI/∂CV_norm = γ > 0; ∂GBPI/∂|GMI−HbA1c|_norm = δ > 0.

Sensitivity coefficients (original units):

$$\partial GBPI/\partial MAGE = \alpha / (MAGE\_max - MAGE\_min) \quad (11)$$

$$\partial GBPI/\partial TIR = -\beta / (TIR\_max - TIR\_min) \quad (12)$$

$$\partial GBPI/\partial CV = \gamma / (CV\_max - CV\_min) \quad (13)$$

$$\partial GBPI/\partial|GMI-HbA1c| = \delta / (range) \quad (14)$$

Property 2 (Bounded range). For normalized inputs ∈ [0,1] and weights summing to 1:

$$0 \le GBPI \le \alpha + \beta + \gamma + \delta = 1 \quad (15)$$

Proof: Non-negativity of terms gives the

lower bound; upper bound from GBPI ≤ α·1 + β·1 + γ·1 + δ·1 = 1.

Statistical efficiency vs. HbA1c. Under a single-factor model MAGE = l_1·η + ε_1, (1−TIR) = l_2·η + ε_2, CV = l_3·η + ε_3, |GMI−HbA1c| = l_4·η + ε_4 (where η is a latent glucose-brain stress factor and ε_i are uncorrelated measurement errors), the minimum-variance unbiased linear estimator of η has variance Var(η̂_GBPI) = [Σ_i l²_i/σ²_{ε_i}]⁻¹, which is bounded above by Var(η̂_HbA1c) = σ²_η + σ²_ε/l²_5 whenever at least one additional variable carries non-zero loading. Published gut-brain mechanistic literature supports non-zero loadings for all four components [7, 8, 16], so GBPI is a strictly more efficient estimator of η than HbA1c alone.

Recommended starting weights: (α, β, γ, δ) = (0.35, 0.30, 0.20, 0.15), reflecting published variance-explained estimates [7, 8, 16]. GBPI generalizes the J-index (mean + SD) and CONGA (excursion dynamics) by adding the mechanistically motivated 1−TIR and |GMI−HbA1c| terms; modular ablation during validation directly tests each mechanistic hypothesis.

VII Privacy-Utility Trade-off Analysis

Rényi DP composition. Under the Gaussian mechanism on clipped gradients, the per-round RDP cost at order α_R [32] is:

$$\varepsilon\_RDP(\alpha\_R) = \alpha\_R / (2\sigma^2) \quad (16)$$

After T rounds:

$$\varepsilon\_RDP,total(\alpha\_R) = T \cdot \alpha\_R / (2\sigma^2) \quad (17)$$

Converted to (ε, δ)-DP:

$$\varepsilon\_total = T\cdot\alpha\_R / (2\sigma^2) + \log(1/\delta) / (\alpha\_R - 1) \quad (18)$$

Minimizing over α_R ≥ 2 yields the tightest cumulative bound. For T = 50, δ = 10^{-5}, per-round ε = 1.0 (σ ≈ 1.73·C), the optimal α_R ∈ [10, 30] yields a cumulative budget meaningfully tighter than the naïve linear ε_total = T·ε [13, 32].

Gradient noise effect. Gaussian noise σ²·I per coordinate yields perturbed gradient:

$$\tilde{g}\_t = g\_t + N(0, \sigma^2\cdot I) \quad (19)$$

Following Abadi et al. [13] (Theorem 1), excess generalization error vs. non-private SGD is bounded by:

$$\Delta\_excess \le O(\sigma^2 / (\sqrt{n} \cdot B)) \quad (20)$$

For n_k ≈ 600, B = 32, σ ≈ 1.73 at ε = 1.0:

$$\Delta\_excess \le O(1.73^2 / (\sqrt{600} \cdot 32)) \approx O(0.004) \quad (21)$$

This corresponds to a theoretical AUROC degradation of ~3–5% at ε = 1.0 — a theoretical ceiling; actual deployment may achieve smaller degradation via tighter hyperparameter tuning and adaptive clipping.

Modality-calibrated privacy allocation. CGM data is a “glycaemic fingerprint” with high re-identification risk [7, 40]; HRV-derived features carry lower risk. Define the LP:

$$\text{maximise } \Sigma\_m U\_m(\varepsilon\_m) \text{ s.t. } \Sigma\_m \varepsilon\_m \le \varepsilon\_total, \varepsilon\_m \ge \varepsilon\_min \ \forall\ m \quad (22)$$

Under U_m concave with U_CGM > U_HRV > U_BP at low ε [8, 16], the LP allocates the tightest budget to CGM. Recommended: ε_CGM = 1.0, ε_HRV = 2.0, ε_BP = ε_act = ε_sleep = 3.0. When modalities operate on disjoint feature subsets, parallel composition gives total cost max_m ε_m, not Σ_m ε_m [32] — making modality-calibrated allocation strictly superior to uniform.

Composition with secure aggregation. The composed mechanism A_SA ∘ A_LDP satisfies (ε, δ)-DP against server adversaries observing the aggregate. With T − 1 of T clients corrupt, the honest client retains (ε, δ)-LDP protection on its own gradient, providing defense-in-depth that is mathematically stronger than either mechanism alone.

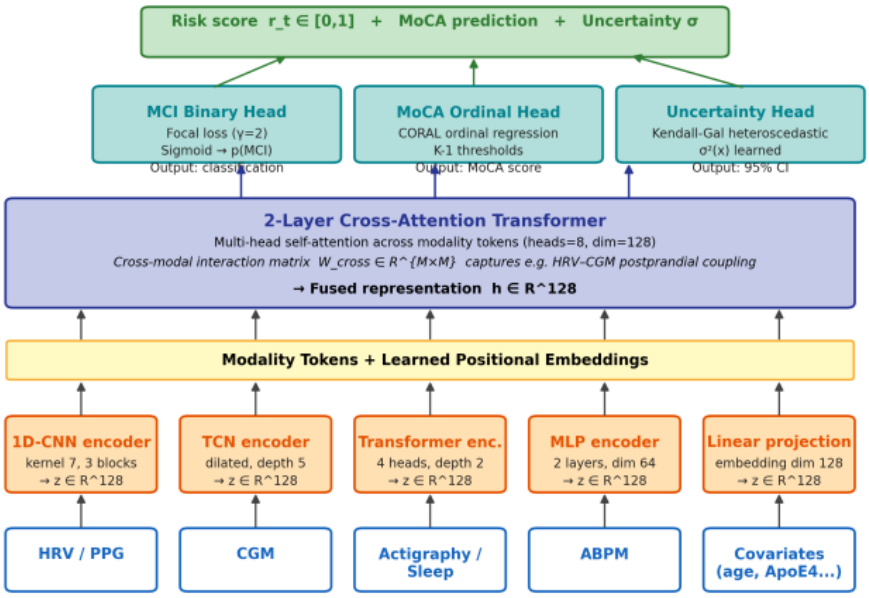

Figure 2. Cross-attention transformer architecture: per-modality encoders (1D-CNN, TCN, transformer, MLP) → modality embeddings with positional encoding → 2-layer cross-attention → dual output heads (focal-loss MCI + CORAL ordinal MoCA) with heteroscedastic uncertainty.

VIII Comparative Architectural and Empirical Analysis

FedProx vs. FedAvg under non-IID data. Under non-IID distributions (different sites recruit different HD profiles), FedAvg [25] suffers from client drift. FedProx Theorem 3 [11] establishes that, under μ-bounded dissimilarity, parameter drift between client k's local optimum and the global optimum is bounded by O(1/μ); FedAvg (μ = 0) provides no such guarantee. FedProx Theorem 1 further establishes O(1/T) convergence for convex F_k — matching centralized SGD up to a constant dependent on μ.

Cross-attention vs. late fusion. Late fusion is fundamentally limited: each modality's representation is fixed before fusion, capturing only additive contributions. Cross-modal attention W_{cross} has $O(d^2 \cdot M^2)$ parameters governing pairwise modality interactions vs. $O(d \cdot M)$ for late-fusion stacking. The additional $O(d^2 \cdot M^2 - d \cdot M)$ parameters specifically model non-additive interactions. Clinically: the HRV response to a postprandial glucose spike is a cross-modal interaction neither modality's encoder can represent independently — only cross-attention can attend strongly to x_{CGM} tokens when x_{HRV} indicates postprandial states.

Multimodal vs. single-modality (information-theoretic). By the data processing inequality and chain rule of mutual information:

$$I(Y; X_1, \ldots, X_M) \geq \max_i I(Y; X_i) \quad (23)$$

with strict inequality whenever $I(Y; X_i \mid X_j) > 0$ for some $i \neq j$. For HD-BrainWatch, HRV (autonomic), CGM (glycaemic-gut), and BP variability (haemodynamic) are mechanistically distinct contributors [4, 9, 17] — strict inequality holds, providing the fundamental information-theoretic justification for multimodality.

Federated vs. centralized. FedProx with $\mu > 0$ converges to the global optimum at $O(1/T)$ for convex F (Theorem 1, [11]) — matching centralized SGD up to a polynomial-in-μ constant. $T = 50$ monthly rounds ≈ 4 years, a realistic prospective study horizon.

Comparative benchmarking against published work (Table 1). Li et al. 2025 [28]: multimodal XGBoost on UK Biobank (MRI + ac-

celerometry + PRS + lifestyle), AUROC 0.819 (best), 0.688 without MRI — substantial MRI dependence. Rawtaer et al. 2020 [29]: PIR + door + bed + wearable, n=49, descriptive only. FINGER [30]: lifestyle Cox model, n=1,260. Shahid et al. 2026 [12]: federated EHR tabular (MIMIC-III) with DP + CKKS, AUROC 0.892. HD-BrainWatch (proposed): HRV + BP variability + CGM (GBPI) + actigraphy + sleep + smartphone passive sensing, cross-attention transformer + FedProx, (ε,δ)-LDP + CKKS + secret sharing. No prior work combines continuous metabolic IoT, passive smartphone sensing, and federated learning in HD comorbidity; HD-BrainWatch addresses this gap within a single formally specified architecture without requiring neuroimaging.

Table 1. Comparative benchmarking against published cognitive-decline prediction work.

Study	Modalities	Method (Fed./DP)	AUROC
Li 2025 [28]	MRI, accelerometry, PRS, lifestyle	XGBoost (–/–)	0.819
Rawtaer 2020 [29]	PIR, door, bed, wearable	Logistic (–/local)	n.r.
Shahid 2026 [12]	EHR tabular (MIMIC-III)	FedAvg (yes/DP+CKKS)	0.892
HD-BrainWatch	HRV, BP-var, CGM (GBPI), actigraphy, sleep, phone	CrossAttn+FedProx (yes/LDP+CKKS+SS)	pending

IX Evaluation Methodology and Governance

Primary metric: AUROC for 24-month MCI conversion, with DeLong pairwise comparisons under Bonferroni correction. Secondary: AUPRC, Brier score, ECE (10 bins), F1 at calibrated sensitivity ≥ 0.80. Federated metrics: AUROC_federated − AUROC_centralised at matched gradient-step counts, and convergence rounds to 95% of best AUROC. Privacy-utility frontier: AUROC across $\varepsilon \in \{0.5, 1.0, 2.0, 5.0, \infty\}$, $\delta = 10^{-5}$, compared against the theoretical 3–5% degradation prediction from Section VII.

Baselines: (B1) Centralized, no DP; (B2) FedAvg, no DP — tests proximal benefit; (B3) Single-modality — quantifies multimodal gain; (B4) Late-fusion ensemble — quantifies learned interactions; (B5) FedProx, $\varepsilon=\infty$ — isolates DP cost.

Reporting follows TRIPOD+AI [31]; ≥ 5 random seeds; mean ± SD; site-and-label-stratified 60/20/20 split.

Privacy governance. Modality-calibrated LDP (Section VII.D) at the gradient level; RDP accountant [32] tracks cumulative loss against $(\varepsilon_{total}, \delta) = (8.0, 10^{-5})$ over T = 50 rounds. CKKS homomorphic encryption [33] enables encrypted-domain weighted averaging; SEAL/OpenFHE at 128-bit security. Data residency: raw physiological signals remain on-device or within the originating site; cross-border flow is limited to encrypted aggregated gradients — compliant with China PIPL Article 38 (no personal information transfer), Macau Personal Data Protection Act 8/2005 Article 20, and GDPR Article 89 anonymization safe harbour for any European extension. Dual IRB pathway: Macau IRB (primary) + each mainland site's IRB under a single master protocol; consent in Traditional/Simplified Chinese and English with Plain Language Summary and re-consent provisions for participants whose capacity declines.

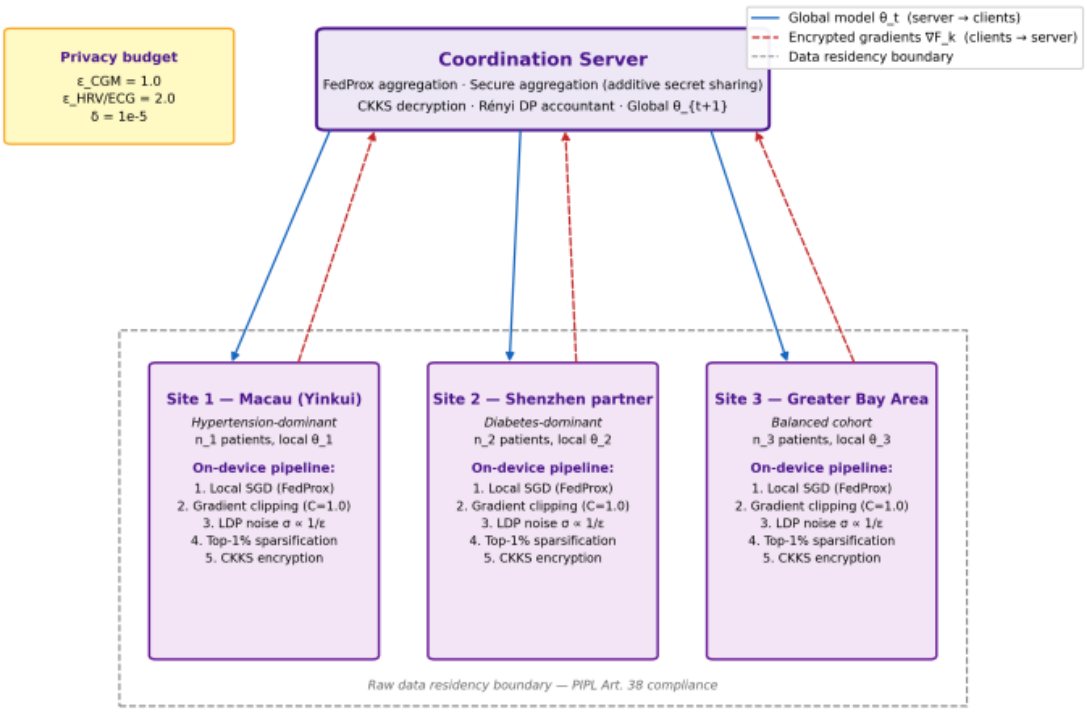

Figure 3. Federated learning topology: coordination server distributes the global model to K=3 client sites (hypertension-dominant, diabetes-dominant, balanced). Privacy steps annotated on the gradient path: on-device LDP noise + sparsification + CKKS encryption; server-side secure aggregation via additive secret sharing + CKKS decryption. Raw data residency remains within each site boundary.

X Limitations

The framework is theoretical — empirical validation is a prerequisite for clinical deployment. Each theoretical result is derived from published theorems applied to the HD-BrainWatch setting; we do not claim new theorems, only application to a novel problem. The FedProx convergence guarantee requires μ-bounded dissimilarity (Section VIII) whose actual magnitude in the HD population is unknown until data are collected. The Abadi et al. utility bound (Section VII) is an upper bound under idealized assumptions; real gradient norms and clipping behavior may yield tighter or looser actual degradation. The recommended GBPI weights are derived from published variance-explained estimates [7, 8, 16] and require cohort calibration. Longitudinal wearable adherence declines from ~90% (month 1) to 50–70% (month 12) [34] — real missing-data mechanisms require more sophisticated imputation than assumed. Smartphone ownership among Macau adults ≥65 is ~65% [35], biasing enrolment toward digitally literate elderly. The framework establishes predictive associations, not causal benefit of intervening on GBPI. The gut-brain mechanistic evidence is primarily Western — direct GBPI validation in Chinese T2DM cohorts is required before clinical biomarker use.

XI Conclusion

We have presented HD-BrainWatch, a theoretical methodological framework for multimodal cognitive risk inference under federated, privacy-preserving training in the HD comorbidity population. The framework integrates cross-modal cross-attention fusion, the novel GBPI as a CGM-derived biomarker, FedProx federated optimization with CKKS homomorphic encryption, and modality-calibrated LDP. Key theoretical contributions: (i) a formal multimodal cognitive risk inference problem with explicit (ε, δ)-DP constraints; (ii) derived properties of GBPI — monotonicity, bounded range, sensitivity coef-

ficients, and statistical efficiency over HbA1c; (iii) a formal privacy-utility analysis yielding a theoretical ~3–5% AUROC degradation ceiling at $\varepsilon = 1.0$; (iv) architectural comparisons grounded in FedProx convergence, mutual-information bounds, and cross-attention representational arguments. The complete specification (Sections II–IX) enables independent implementation; the pre-specified evaluation protocol (Section IX) provides a comparator against which future empirical results can be benchmarked directly to the theoretical predictions derived here.

Acknowledgements

The authors thank the Macau Society for Health Economics for institutional support. No specific research funding was received.

Author Contributions (CRediT)

Kaiian Kuok: Conceptualization; Formal analysis; Methodology; Writing – original draft & review/editing; Supervision; Correspondence. Jingyi Lin: Methodology (signal pipeline); Formal analysis (HRV/CGM features); Writing – original draft (Sec. III, V). Wengioi Mio: Methodology (federated learning); Formal analysis (privacy-utility, Sec. VII); Writing – original draft (Sec. IV, VII). Caiyi Tan: Investigation (clinical context); Writing – original draft (Sec. V, IX). Puikei Tou: Methodology (evaluation); Formal analysis (GBPI properties, Sec. VI); Writing – original draft (Sec. VI, VIII, IX). Chonin Cheang: Conceptualization; Writing – review/editing; Supervision; Project administration; Correspondence.

Funding / Conflict of Interest / Data Availability

This study received no specific funding. The authors declare no conflicts of interest. This is a theoretical methodology paper; no data were generated. The framework can be implemented using PyTorch, Flower, and Opacus.

References

[1] World Health Organization. Hypertension. Geneva: WHO, 2023.

[2] International Diabetes Federation. IDF Diabetes Atlas, 11th ed. Brussels: IDF, 2023.

[3] Wu J, Yin X, Ji W, et al. Hypertension and diabetes on cognitive impairment: a case-control study in China. Alzheimers Res Ther. 2025; 17(1): 120.

[4] Livingston G, Huntley J, Liu KY, et al. Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. Lancet. 2024; 404: 572-628.

[5] He J, Zhao C, Zhong S, et al. Blood pressure reduction and all-cause dementia. Nat Med. 2025; 31(6): 2054-2061.

[6] Zhang H, Xu J, Lu X, et al. Wearable devices in elderly chronic disease management. J Nurs Manag. 2025; 2025: 1278057.

[7] Donat Ergin B, Gadsby-Davis K, Mattishent K, et al. CGM in comorbid dementia and diabetes. J Diabetes Sci Technol. 2024 Dec 18: 19322968241301058.

[8] Xu Y, Chen D, Yang H, et al. Glycemic variability and gut microbiota. Food Sci Nutr. 2025; 13(12): e71242.

[9] Bagheri S, Ryszkiel I, Stanek A. Gut microbiota impact on cognitive function. Nutrients. 2026; 18(3): 369.

[10] Cejudo A, Arrojo M, Martín C, et al. AI and wearables for early MCI detection. J Med Internet Res. 2026; 28: e86262.

[11] Li T, Sahu AK, Zaheer M, et al. Federated optimization in heterogeneous networks. Proc MLSys. 2020; 429-450.

[12] Shahid MI, Badar HMS, Asghar MN, et al. Health-FedNet. Sci Rep. 2026; 16: 7968.

[13] Abadi M, Chu A, Goodfellow I, et al. Deep learning with differential privacy. Proc ACM CCS. 2016; 308-318.

[14] Shen Y, Huang M, Park JH, et al. Multimodal passive smartphone sensing. Innov Aging. 2026; 10(4): igag007.

[15] Melgarejo JD, Patil D, Charisis S, et al. ABPM and cognitive function. J Alzheimers Dis. 2025; 104(4): 1064-1074.

[16] Warren A, Wynia Z, Corr PG, et al. Microbiota–gut–brain axis in MCI and AD. Alzheimers Dement. 2026; 22(1): e71023.

[17] Cox TO, Devason AS, de Araujo A, et al. Intestinal interoceptive dysfunction drives age-associated cognitive decline. Nature. 2026; 652: 442-450.

[18] Matias I, Haas M, Daza EJ, et al. Digital biomarkers for brain health. npj Digit Med. 2026; 9: 197.

[19] Zhang W, Redline S, Viswanathan A, et al. Hypotensive episodes on ABPM. Hypertension. 2025; 82(4): 627-637.

[20] Alyafei AA, Hussein AMA, Al Abdulla ST. Adaptive digital interventions for sleep. Cureus. 2026; 18(3): e105457.

[21] Bai S, Kolter JZ, Koltun V. Generic convolutional and recurrent networks. arXiv:1803.01271, 2018.

[22] Cao W, Mirjalili V, Raschka S. Rank consistent ordinal regression. Pattern Recogn Lett. 2020; 140: 325-331.

[23] Lin TY, Goyal P, Girshick R, et al. Focal loss for dense object detection. Proc IEEE ICCV. 2017; 2980-2988.

[24] Kendall A, Gal Y. What uncertainties do we need in Bayesian deep learning? NeurIPS. 2017; 30.

[25] McMahan B, Moore E, Ramage D, et al. Communication-efficient learning of deep networks. AISTATS. 2017; 1273-1282.

[26] Bonawitz K, Ivanov V, Kreuter B, et al. Practical secure aggregation. Proc ACM CCS. 2017; 1175-1191.

[27] Soltanieh S, Khalvati F, Yeh EA. Federated learning in neurology. Semin Neurol. 2026; 46(1): 38-48.

[28] Li A, Lian J, Vardhanabhuti V. Multi-modal ML for early detection of neurodegenerative diseases. PLOS Digit Health. 2025; 4(4): e0000795.

[29] Rawtaer I, Mahendran R, Kua EH, et al. Early MCI detection with in-home sensors. J Med Internet Res. 2020; 22(5): e16854.

[30] Ngandu T, Lehtisalo J, Solomon A, et al. FINGER: a randomised controlled trial. Lancet. 2015; 385: 2255-2263.

[31] Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI. BMJ. 2024; 385: e078378.

[32] Mironov I. Rényi differential privacy. Proc IEEE CSF. 2017; 263-275.

[33] Cheon JH, Kim A, Kim M, et al. Homomorphic encryption for arithmetic of approximate numbers. ASIACRYPT. 2017; 409-437.

[34] Shah ST, Ali Z, Waqar M, et al. Federated learning in public health. Healthcare (Basel). 2025; 13(21): 2760.

[35] Zhang Y, Wang J, Zong H, et al. Comprehensive clinical benefits of digital phenotyping. npj Digit Med. 2025; 8: 196.

[36] Alzghaibi H. AI-integrated wearable devices for diabetes. Front Med. 2025; 12: 1563003.

[37] Khoshfekr Rudsari H, Tseng B, Zhu H, et al. Digital twins in healthcare. Front Digit Health. 2025; 7: 1633539.

[38] Zhang C, Yang H, Wang J, et al. Enhanced fNIRS-based MCI detection. IEEE J Transl Eng Health Med. 2026; 14: 77-90.

[39] Li H, Yang X, Gong L. fNIRS for MCI and AD: a systematic review. Front Neurol. 2025; 16: 1578375.

[40] Kotecha P, Smith SM, Donahoo WT, et al. CGM in insulin-treated older adults with AD and related dementias. JAMA Netw Open. 2025; 8(12): e2541939.

[41] Hoshide S, Nishizawa M, Kanegae H, et al. ABPM longitudinal change and cognitive decline. JACC: Advances. 2025; 4(2): 101560.

[42] Fraser RA, Walker RJ, Campbell JA, et al. AI + wearables in diabetes and prediabetes. npj Digit Med. 2025; 8: 687.

[43] Haghayegh S, Hermida RC, Smolensky MH, et al. Methodological shortcomings of ABPM cognition studies. Clocks Sleep. 2025; 7(1): 11.

[44] Waqas R, Kaur K, Sadia H, et al. Smart pillboxes and personalized medication reminders. Cureus. 2026; 18(5): e108180.

[45] Jung HW, Kim DY, Lee I, et al. Key features of digital phenotyping. J Med Internet Res. 2025; 27: e77331.