Joint Progression Modeling (JPM): A Probabilistic Framework for Mixed-Pathology Progression

Hongtao Hao, Joseph L. Austerweil

Introduction

Problem: Event-based models (EBMs) assume single-disease pathology, but 45.8% of clinically probable AD patients have mixed pathologies (e.g., AD + Lewy bodies). Treating mixed-pathology as a new single disease is data-inefficient: NACC has 13,400 AD-only patients but only 50 with AD+VaD and complete biomarkers data necessary AD diagnosis and prognosis.

Our Solution: Integrate knowledge from single-disease cohorts to infer joint progression.

ID	Impacted	ABETA	PTAU	MMSE	WMH	...
1	Yes	846.90	11.53	30.90	0.45	...
2	Yes	1187.20	12.32	29.12	0.52	...
3	No	923.78	21.54	30.05	0.21	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Key Question: If AD progression is A → C → D and VaD progression is C → B, how much more likely is A → C → B → D than D → A → B → C for AD + VaD?

Methods

Joint Progression Modeling Framework

JPM obtains the posterior of aggregate rankings (σ) given mixed-pathology data (D) and partial rankings ($\mathcal{S}$):

$$P(\sigma \mid D, \mathcal{S}) \propto P(D \mid \sigma) \cdot P(\sigma \mid \mathcal{S}) \tag{1}$$

Where $P(D \mid \sigma)$ is the data likelihood which can be obtained via methods such as EBMs.

We adopt an energy-based prior where low energy indicates high probability:

$$P(\sigma \mid \mathcal{S}) \propto \exp(-E(\sigma \mid \mathcal{S})) \tag{2}$$

Four Energy Function Variants

Variant	Energy Function	Key Feature
PP	$E = -\sum_{i,j} w_{ij} \mathbf{1}\left[i \underset{\sigma}{<} j\right]$	Weighted pairwise voting
BT	$E = -\log P\left(\sigma \mid \hat{\theta}\right)$	Latent strength parameters
PL	$E = -\log P(\sigma \mid \hat{\alpha})$	Sequential generation
Mallows	$E = \theta \cdot d(\sigma, \sigma_0)$	Distance from center

Inference: Use Metropolis-Hastings MCMC with SA-EBM likelihood.
Generation: Sample aggregate rankings from partial rankings for synthetic data.

Theoretical Analysis

Q1: Is JPM Valid for Inference? (Calibration)

Definition: Does lower energy predict closer distance to ground truth?

$$E_{\text{JPM}(\sigma_A)} < E_{\text{JPM}(\sigma_B)} \implies d\left(\sigma_A, \sigma_{\text{gt}}\right) < d\left(\sigma_B, \sigma_{\text{gt}}\right) \tag{3}$$

We measure calibration as Spearman's ρ between energies and distances to ground truth.

Gen. Variant	BT	Mallows ($\theta = 1$)	PL	PP
BT	0.80	**0.98**	0.61	0.88
PL	0.85	0.91	0.66	0.84
PP	0.79	0.94	0.59	**0.91**

Table 1: Calibration (Spearman's ρ): All variants > 0.58

Result: All JPM variants are valid inference algorithms.

Q2: Is JPM Valid for Generation?

Separation: Can JPM distinguish true ranking from random noise? In other words, the ground truth aggregate ranking must be more consistent with the evidence from the input partial rankings ($\mathcal{S}$) than a random ranking is. **All variants have perfect separation.**

Sharpness: How stable is output across runs (Kendall's W)?

BT: 0.96 | PL: 0.74 | PP: 0.86 | Mallows($\theta = 1$): 0.47 | Mallows($\theta = 10$): 0.61

Result: Near-perfect separation; sharpness varies by variant.

Q3: Predicting Sharpness

Sharpness is highly predictable ($R^2 > 0.6$ for BT, PL, Mallows) from four characteristics of $\mathcal{S}$:

- Number of partial rankings (K)
- Average length ($\bar{\ell}$)
- **Conflict:** $\frac{2}{K(K-1)} \sum_{i<j} d_\tau\left(\sigma^{(i)}, \sigma^{(j)}\right)$
- **Overlap:** Proportion of biomarkers in ≥ 2 rankings

Usage: Choose desired sharpness → calculate these 4 features → select matching variant.

Experiments & Results

Synthetic Experiments

Setup: 4,050 datasets with 18 ADNI biomarkers, varying:

- Participant size (J): 50, 100, 200
- Healthy ratio (R): 0.25, 0.5, 0.75
- 5 JPM variants × 9 experimental configurations × 10 random datasets per $J - R$ pair (9 pairs).

Evaluation: (1) Ordering accuracy (normalized Kendall's τ distance), (2) Staging accuracy (MAE)

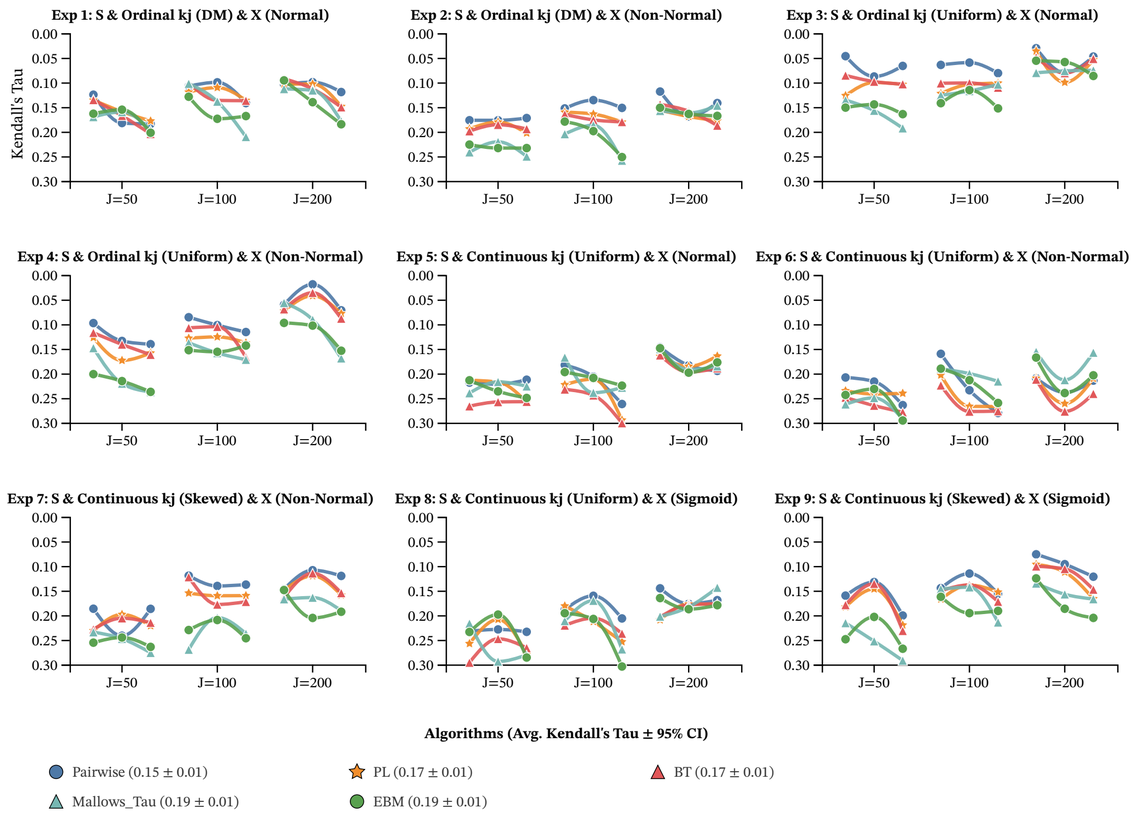

Figure 1: **PP-generated data:** JPM achieves normalized τ distance of 0.15 vs. SA-EBM's 0.19

Key Findings:

- JPM outperforms SA-EBM when data matches high-sharpness variants (PP, BT, PL)
- Advantage greatest with small sample sizes (J=50, 100)
- Even with mismatched variants, JPM remains competitive

Data Gen.	Best JPM	JPM τ	vs. Baseline
BT	BT, PL, PP	**0.15**	21% better
PL	PL	**0.17**	15% better
Mallows	Mallows	**0.19**	Comparable

NACC Real-World Data

Datasets: AD-only (188 healthy, 59 patients), VaD-only (3,732 healthy, 525 patients), AD+VaD (188 healthy, 37 patients)

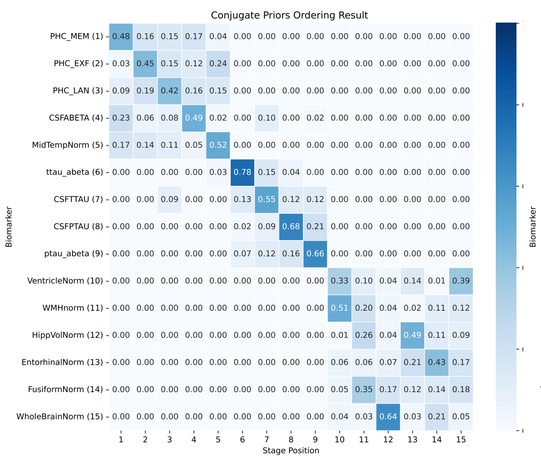

Figure 2: SA-EBM result

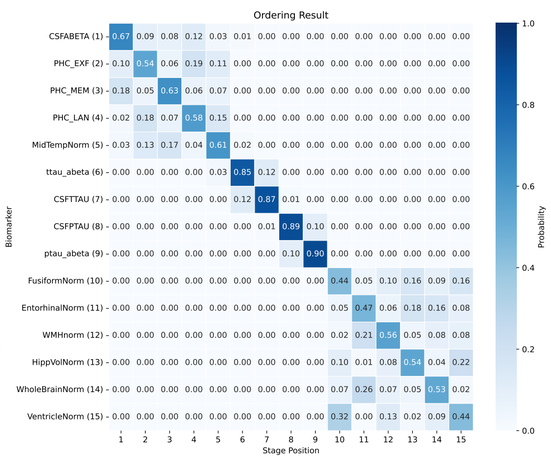

Figure 3: JPM-Mallows result

Finding: Both SA-EBM and JPM-Mallows position vascular damage (WMHnorm) **after** initial molecular pathology but **before** widespread atrophy, supporting synergistic disease interaction. JPM-Mallows slightly better aligns with literature showing amyloid/tau buildup precedes cognitive symptoms.

Discussion & Conclusions

- Introduced JPM: first framework for mixed-pathology progression modeling
- All variants show strong calibration and separation
- 21% improvement over single-disease baseline in synthetic experiments
- Real-world validation: results align with neurodegeneration literature

When to use JPM:

- **High-sharpness scenarios:** JPM >> single-disease model
- **Low-sharpness:** JPM still competitive, reliable
- **Variant selection:** Estimate sharpness from partial rankings, choose matching variant

Limitations & Future Work:

- Assumes shared progression (extend to subtypes like SuStaIn)
- Ordinal events only (integrate Temporal EBM for continuous time)
- Requires domain knowledge of disease interactions

Impact: Enables principled modeling of comorbidities using existing single-disease data, addressing critical gap in neurodegeneration research.

Code: `pip install pyjpm` | **Package:** github.com/jpcca/pyjpm
Data & Experiments: github.com/hongtaoh/jpm
Acknowledgement: ADNI/NACC for data, CHTC for computing resources, and JPCCA for funding

 hongtaoh@cs.wisc.edu