

BACKGROUND

- Harmonized cognitive composites are widely used to study longitudinal cognitive decline across studies.
- Although psychometric harmonization places scores on a common latent scale, measurement precision can still vary substantially across visits, age, and diagnostic groups (Figure 1).
- Standard longitudinal mixed-effects models ignore this heterogeneity, potentially biasing estimates of decline and their clinical interpretation.

METHODS

- Analyzed harmonized memory scores and visit-specific standard errors derived from confirmatory factor analysis of the National Alzheimer’s Coordinating Center Uniform Data Set (Table 1).
- Compared three longitudinal modeling approaches: (1) a naïve mixed-effects model assuming constant measurement precision; and two precision-aware models — (2) a precision-weighted frequentist mixed-effects model that down-weights observations with larger standard errors; and (3) a Bayesian measurement-error mixed-effects model that treats observed scores as noisy realizations of an underlying latent memory trajectory.
- Models adjusted for age, sex, education, APOEε4 status, and diagnostic group (normal cognition [NC], mild cognitive impairment [MCI], Alzheimer’s dementia [AD]), with participant-specific random intercepts and slopes.
- To assess clinical significance, we estimated the age at which predicted memory trajectories crossed a clinically meaningful impairment threshold, defined as 1 SD below the mean memory score for NC individuals at age 60.

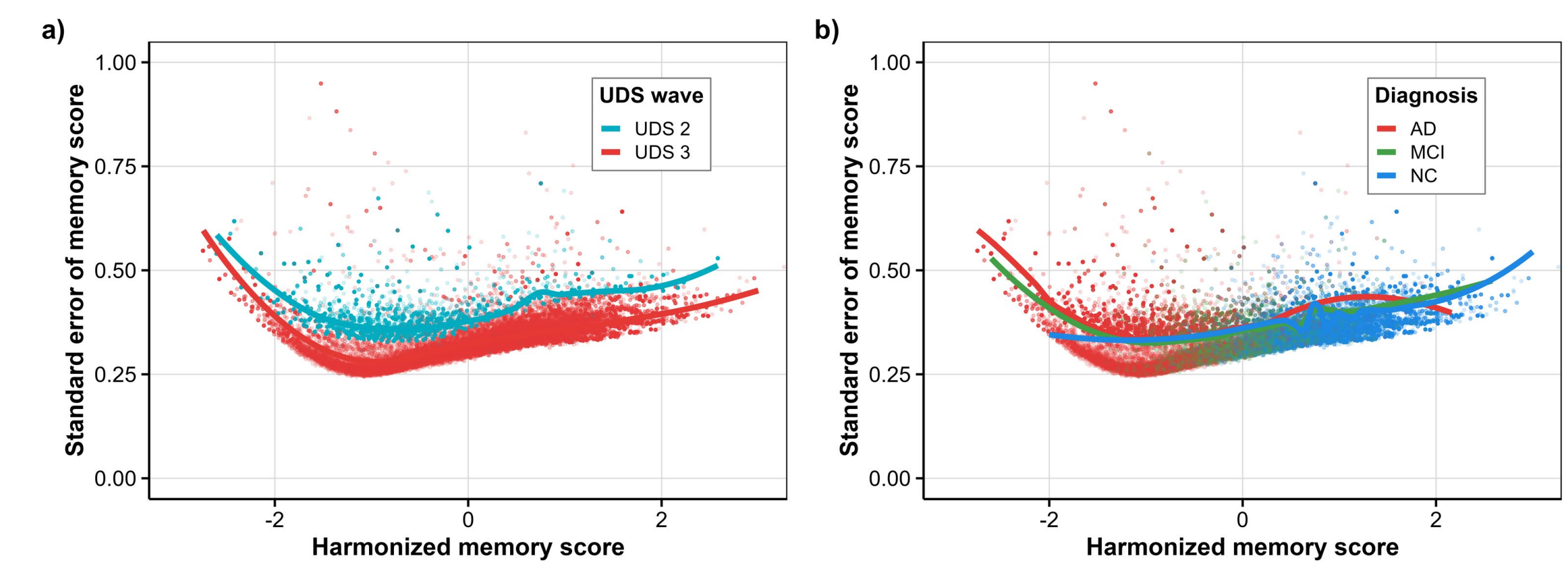
TABLE 1

Baseline demographic and clinical characteristics. Baseline = first visit with age ≥ 60. NC = normal cognition; MCI = mild cognitive impairment; AD = Alzheimer’s dementia.

Characteristic	Overall	NC	MCI	AD
N	25,898	12,964	6,307	6,627
Age at baseline, mean (SD)	73 (7.9)	72.2 (7.8)	73.8 (7.6)	74.1 (8.3)
Female, %	57	65	48	49
Education, yrs, mean (SD)	15.6 (3.2)	16 (2.9)	15.5 (3.2)	14.8 (3.5)
APOE ε4 carrier, %	40	31	44	55
Visits, median (IQR)	4 (2–6)	5 (3–8)	3 (2–5)	3 (2–4)

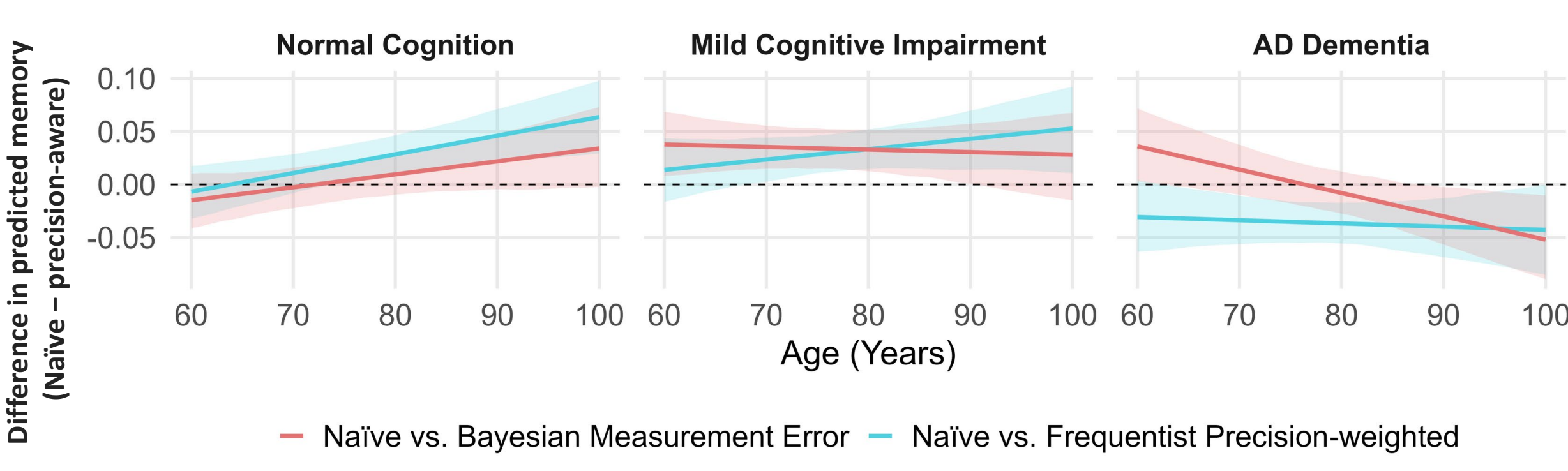
1 Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine, Seattle, WA, USA 2 VA Puget Sound Health Care System, GRECC, Seattle, WA, USA 3 Department of Medicine, University of Washington, Seattle, WA, USA 4 Department of Neurology, Boston University Chobanian & Avedisian School of Medicine, Boston, MA, USA 5 Indiana Alzheimer’s Disease Center and Center for Neuroimaging, Indiana University School of Medicine, Indianapolis, IN, USA 6 Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA

FIGURE 1



Visit-specific standard error of the harmonized memory score varies systematically with ability, UDS wave (a), and baseline diagnosis (b). Standard errors are smallest mid-range and inflate toward the floor and ceiling — the empirical basis for precision-aware modeling.

FIGURE 2



Age-dependent bias in memory trajectories when measurement precision is ignored, stratified by baseline diagnosis. The y-axis is the difference (naïve – precision-aware prediction) by age. Positive values: the naïve model predicts higher memory, reflecting attenuation of estimated decline. Negative values: the naïve model predicts lower memory. The dashed line denotes no bias. Lines compare the naïve model to the frequentist precision-weighted and the Bayesian measurement-error (ME) models; shaded bands are 95% intervals. All predictions are marginal (population-average) estimates.

CONCLUSIONS

- Even with harmonized cognitive scores, ignoring visit-specific measurement precision biases longitudinal inference.
- Accounting for measurement uncertainty yields more realistic trajectories and advances the estimated onset of clinically meaningful memory impairment by years.
- To improve statistical efficiency and power to detect cognitive change, precision-aware modeling should be considered in longitudinal cognitive-aging analyses and in the design and implementation of clinical trials.

RESULTS

- Across diagnostic groups, accounting for measurement precision systematically altered estimated memory trajectories relative to the naïve model.
 - In NC and MCI participants, the naïve model consistently overestimated memory at older ages, which attenuates age-related decline estimates; differences between naïve and precision-aware models increased with age (Figure 2).
 - In contrast, among individuals with AD, the naïve model underestimated memory at older ages due to high variability and floor effects.
 - Only modest differences found between precision-aware approaches.
- In time-to-threshold analyses, the naïve model predicted clinically meaningful memory impairment at 89.2 years, whereas precision-aware models estimated threshold crossing 1.5–3 years earlier (86.2 years frequentist; 87.7 years Bayesian).

TABLE 2

Predicted population-average memory score (naïve model) and the bias from ignoring measurement precision (naïve – precision-aware), shown at the youngest and oldest modeled ages. Bias is small at 70 and accumulates with age (see Figure 2). PW = precision-weighted; ME = measurement error.

Diagnosis	Age	Naïve	Frequentist	Bayesian
NC	70	0.912	0.011	0.006
	100	0.518	0.064	0.043
MCI	70	0.496	0.024	0.045
	100	-0.033	0.053	0.037
AD	70	-0.241	-0.034	0.023
	100	-0.880	-0.043	-0.043