Diabetes and Cognitive Decline: Causal Evidence under Positivity Violations

Zihan Zhang [1] Bryan Keller [2]

[1]Department of Statistics, Columbia University [2]Teachers College, Columbia University

zz3213@columbia.edu
(984) 218-7517

Background

Diabetes is highly prevalent among older adults and has long been linked to accelerated cognitive decline.
However, identifying its **causal effect** is challenging because diabetes onset is time-varying and highly correlated with health status, healthcare use, and survival.
Standard regression and marginal structural models often suffer from **extreme positivity violations**, making the "ever vs. never diabetes" contrast largely non-identifiable.
This study uses longitudinal data from the Health and Retirement Study and a design-based approach to estimate the cognitive effect of incident diabetes among comparable older adults.

Research Question

Does **incident diabetes (0→1)** cause a change in **cognitive function** six years later among older adults in the Health and Retirement Study?

Data & Variables

Dataset. We use the Health and Retirement Study (HRS), a nationally representative cohort of U.S. adults aged 50+, interviewed biennially from 1992–2020. All waves were first assembled into a person–wave panel for marginal structural models.

Initial MSM Dataset. The full longitudinal panel included:

- time-varying diabetes status,
- time-varying health, behavioral, socioeconomic covariates,
- repeated cognition scores.

MSM-based IPTW for the "ever vs. never" contrast showed severe **positivity violations**.

Risk-Set Dataset (Final Analysis). To obtain an identifiable contrast, we restricted to respondents who:

- had no diabetes at Wave 7, and
- had diabetes status observed at Wave 8.

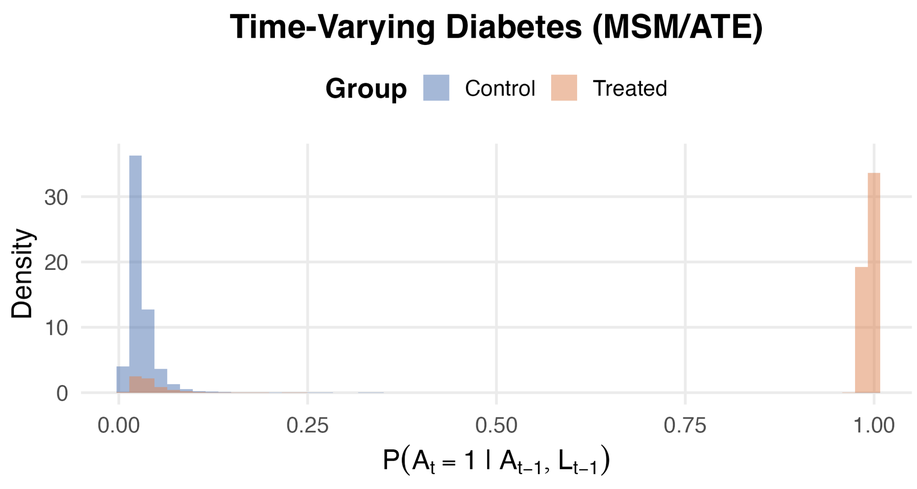
Incident diabetes = transition 0→1 (W7–W8). Outcome = cognition at Wave 10 (6-year follow-up).

Measures.

- **Exposure:** incident diabetes (0→1).
- **Outcome:** cognition score (0–35).
- **Baseline covariates:** demographics; BMI, hypertension, heart disease, stroke; smoking, drinking, walking, sleep; income, work; baseline cognition.

Initial Check: MSM + IPTW for Ever vs. Never Diabetes

We first attempted to estimate the average treatment effect of *ever* versus *never* having diabetes using a marginal structural model with stabilized inverse probability of treatment weights, based on the full longitudinal panel.



MSM/IPTW propensity score distribution.

Figure. Estimated propensity scores for the "ever diabetes" MSM. Non-diabetic individuals have scores concentrated near 0, whereas those who ever develop diabetes have scores near 1, leaving almost no common support. This indicates severe **positivity violation**, making the ATE for ever vs. never diabetes non-identifiable under standard IPTW.

Study Design

Motivation. The MSM/IPTW analysis for the ever vs. never diabetes contrast showed extreme positivity violations. This made the ATE non-identifiable.

Risk-Set Construction. To obtain an identifiable causal contrast, we focused on incident diabetes within a well-defined risk set. Respondents were included if they:

- had no diabetes at Wave 7 (baseline), and
- had observed diabetes status at Wave 8.

These individuals form a risk set where an **incident diabetes** transition is defined as:

$$0 \to 1 \text{ between Waves 7–8.}$$

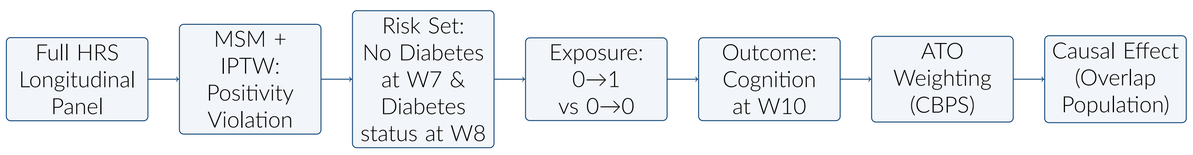
Cognition is measured 6 years later at Wave 10.

Target Estimand. Within the risk set, we estimate the causal contrast:

$$\mathbb{E}\big[Y(0 \to 1) - Y(0 \to 0)\big],$$

representing the effect of **incident diabetes** among individuals who were comparably at risk of developing diabetes.

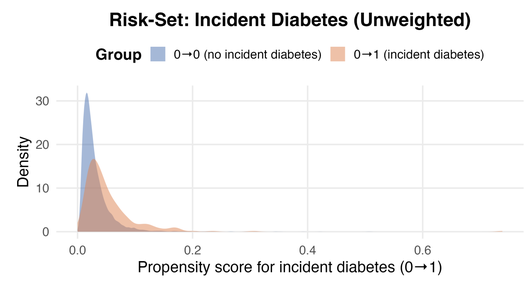
Why This Restores Identifiability. Focusing on incident transitions avoids lifetime exposure separation. Overlap weighting (ATO) further reduces poorly comparable individuals, restoring meaningful common support.



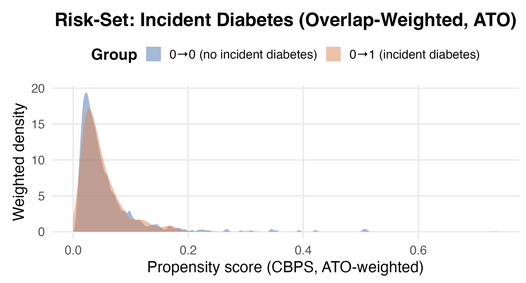
Assumption Check

1. Positivity.

We evaluated whether individuals with incident diabetes (0→1) and those who remained non-diabetic (0→0) had comparable probabilities of developing diabetes, conditional on baseline covariates.



Unweighted propensity score overlap.



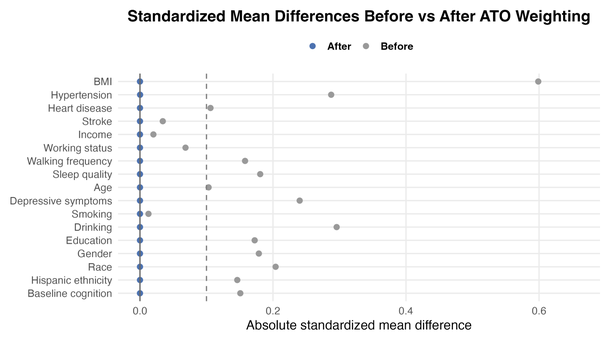
ATO-weighted propensity score overlap.

2. Covariate Balance.

Standardized mean differences (SMDs) were examined to assess covariate balance before and after weighting.

- **Before weighting:** several covariates showed moderate imbalance (>0.1).
- **After ATO weighting:** all covariates achieved excellent balance (SMD ≈ 0).

Interpretation of the ATO population. ATO weighting focuses inference on individuals with moderate diabetes risk, reducing the weight of those at extremely low or high risk and producing a more comparable overlap population.



Covariate balance before vs. after ATO weighting.

Statistical Model, Results, & Interpretation

Statistical Model. Within the risk-set population, we estimate the causal effect of incident diabetes on 6-year cognition using a weighted linear model:

$$Y_{i,\,6\text{yr}} = \beta_0 + \beta_1 D_i + \varepsilon_i,$$

where

- $D_i = 1$ if respondent transitioned 0→1 (incident diabetes) between Waves 7–8;
- $Y_{i,\,6\text{yr}}$ is the global cognition score at Wave 10;
- weights = overlap weights (ATO) from CBPS, defining the overlap population of comparable individuals.

We report two estimates for comparison:

1. **Unweighted model** — naive comparison (confounding not addressed).
2. **ATO-weighted model** — causal contrast in the overlap population.

Results. Table displays regression estimates and 95% confidence intervals:

Model	Estimate	95% CI
Unweighted	$\hat{\beta}_1 = 0.46$	[−0.43, 1.34]
ATO-weighted (CBPS)	$\hat{\beta}_1 = 1.15$	[0.37, 1.93]

The weighted estimate corresponds to the causal contrast

$$\mathbb{E}[Y(0 \to 1) - Y(0 \to 0)]_{\text{ATO}},$$

representing the effect of incident diabetes among individuals who were comparably at risk of developing diabetes.

Interpretation. The unweighted model shows no clear causal effect between incident diabetes and later cognition. After applying overlap weighting, which creates a balanced comparison group, incident diabetes causally lowers 6-year cognition by **1.15-point lower** (95% CI: 0.37–1.93). This suggests a meaningful negative cognitive impact of new-onset diabetes among older adults who were comparably at risk.

Conclusion & Limitations

Conclusion

- MSM/IPTW estimation for the ever/never diabetes contrast failed due to severe positivity violations.
- A risk-set design targeting **incident diabetes** yields an identifiable and meaningful causal contrast.
- Overlap weighting (ATO) restores covariate balance and restricts inference to comparable individuals.
- Incident diabetes causally lowers 6-year cognition by about 1.15 points in the overlap population.

Limitations

- Diabetes and covariates rely on self-reported measures, which may introduce misclassification.
- Overlap weighting balances observed covariates but does not eliminate potential unmeasured confounding.

References

- Zhu, Y., Hubbard, R. A., Chubak, J., Roy, J., & Mitra, N. *Core concepts in pharmacoepidemiology: violations of the positivity assumption in the causal analysis of observational data*. Pharmacoepidemiology and Drug Safety, 2021.
- Cheng, G., Huang, C., Deng, H., & Wang, H. *Diabetes as a risk factor for dementia and mild cognitive impairment: a meta-analysis of longitudinal studies*. Journal of Alzheimer's Disease, 2012.
- Qi, X., Pei, Y., Malone, S. K., & Wu, B. *Social isolation, sleep disturbance, and cognitive functioning (HRS): A longitudinal mediation study*. Journal of Gerontology: Psychological Sciences, 2023.
- Xue, M., Xu, W., Ou, Y.-N., Cao, X.-P., Tan, M.-S., Tan, L., & Yu, J.-T. *Diabetes mellitus and risks of cognitive impairment and dementia: a systematic review and meta-analysis of 144 prospective studies*. Ageing Research Reviews, 2019.
- Health and Retirement Study (HRS). Public use dataset. Institute for Social Research, University of Michigan. `https://hrsdata.isr.umich.edu/data-products/public-survey-data`