Predicting Adolescent Changes in Cortical Thickness from a Single Timepoint

Zahra Moradimanesh[1], Felix Hoffstaedter[1,2], Kaustubh R. Patil[1,2], Federico Raimondo[1,2], Simon B. Eickhoff[1,2], Jürgen Dukart[1,2], Casey Paquola[1]

1 Institute for Neuroscience and Medicine (INM-7), Research Center Jülich
2 Institute for Systems Neuroscience, Heinrich Heine University Düsseldorf

Multiscale **Ne**urodevelopment Lab
Member of the Helmholtz Association

2652 OHBM 2026
JÜLICH Forschungszentrum

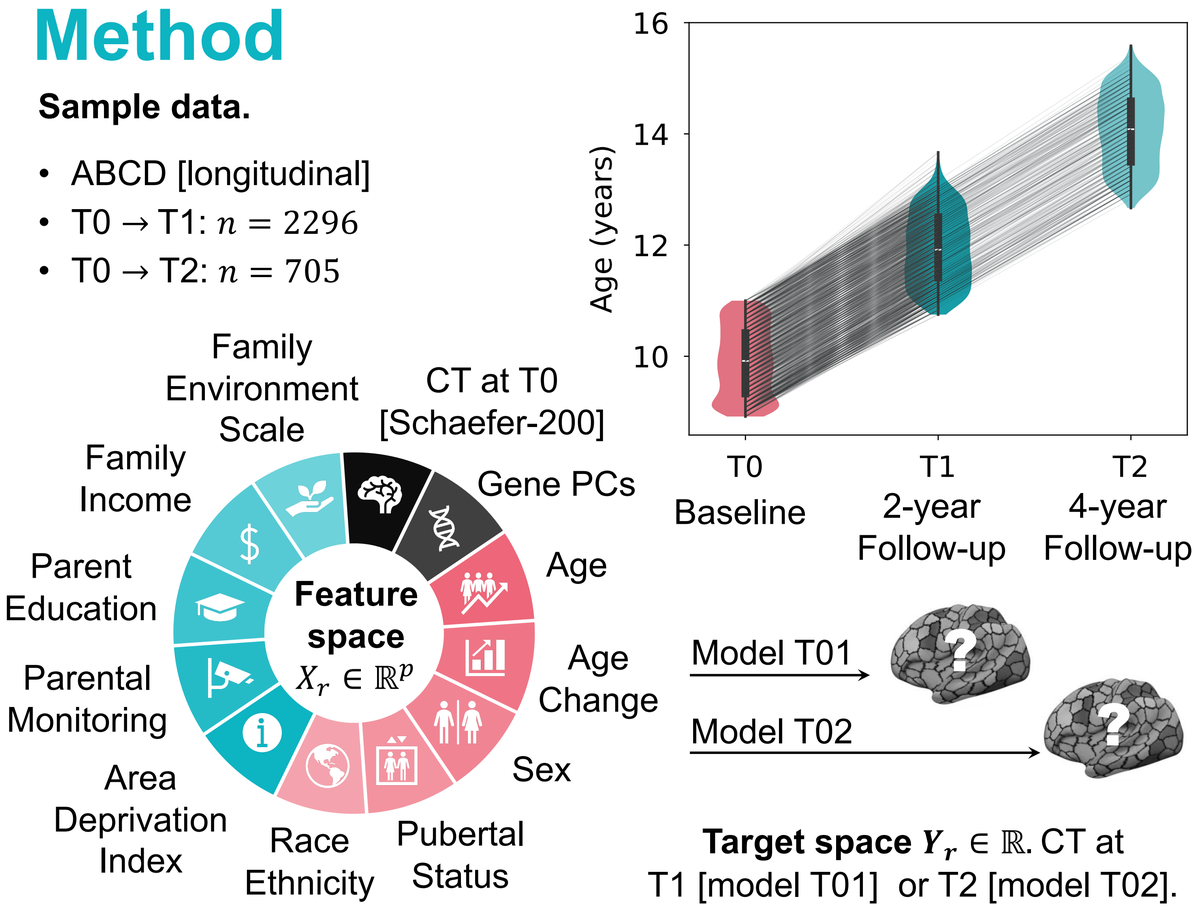
Introduction

- Neurodevelopment lays the foundation for lifelong brain health. Deviations therein underpin various neurological and psychiatric disorders. Identifying and characterizing these deviations early is a central goal of clinical neuroscience.
- It is well-established that intra-individual changes are a stronger predictor of behavioral and clinical outcomes, compared to inter-individual differences[1]. However, it is impractical to rely on multi-year longitudinal assessments to gauge an individual's neurodevelopment, and effective prognosticating depends on predicting changes yet to come.
- We propose a novel solution to infer neurodevelopmental trajectories from a single timepoint by framing the challenge as a longitudinal prediction problem, here focusing on cortical thickness (CT).

Method

Sample data.

- ABCD [longitudinal]
- T0 → T1: $n = 2296$
- T0 → T2: $n = 705$



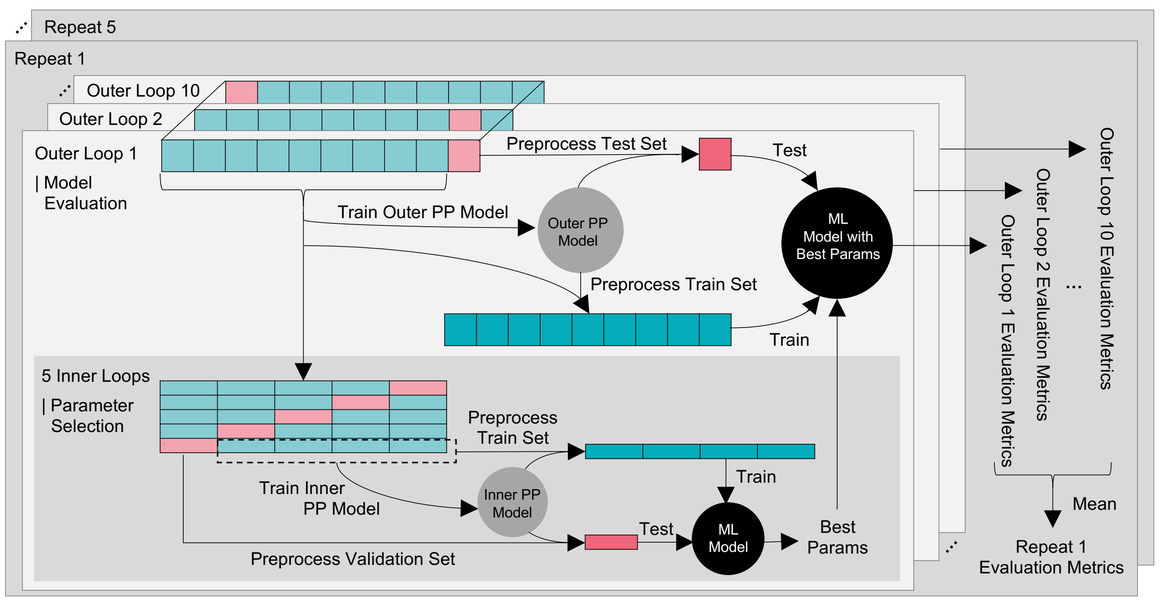
Target space $Y_r \in \mathbb{R}$. CT at T1 [model T01] or T2 [model T02].

Task definition. Prediction of CT at T1 [model T01] and T2 [model T02] from the feature space at T0 is formulated as a supervised conditional distribution estimation problem. For each of the 200 cortical regions, rather than modeling only the conditional mean, we estimate the conditional distribution $F_{Y_r|X_r}(y|x)$, by predicting a set of conditional quantiles $\{Q_{Y_r}(\tau_k\,|X_r)\}_k^K$, $K = 25$ levels spanning uniformly in $[0.02, 0.98]$.

Proposed model. A multi-quantile regression neural network [QRNN], that jointly estimates all conditional quantiles by optimizing a composite pinball loss, yielding the full conditional distribution in a single model while avoiding quantile crossing[2, 3].

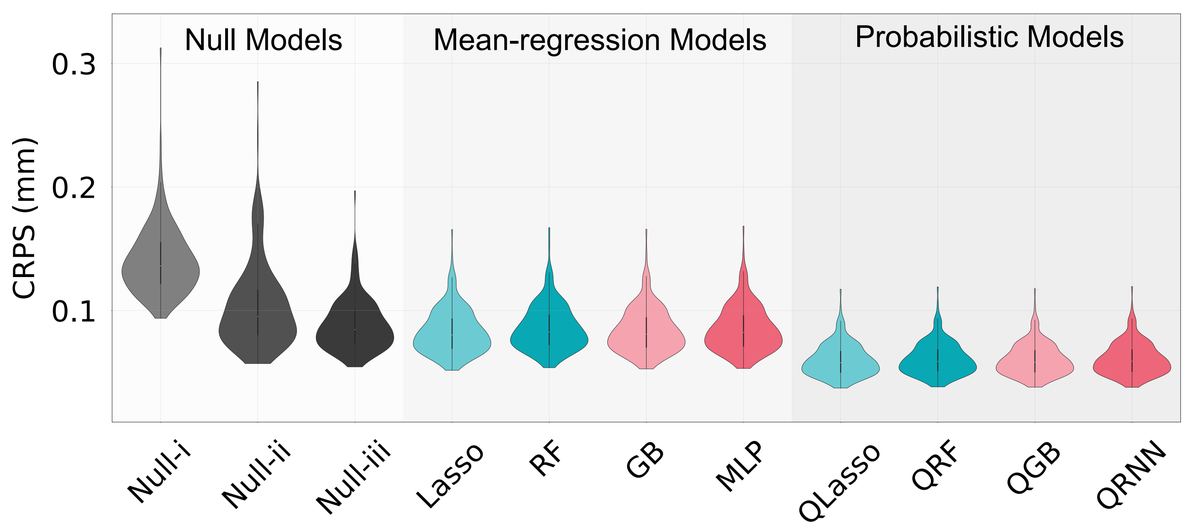
Null and benchmark models. Three parameter-free null models defined as, Null-i [follow-up = group-average baseline], Null-ii [follow-up = individual baseline], and Null-iii [follow-up = individual baseline + group-average change]. In addition, benchmarks comprised deterministic models, estimating the conditional mean: Lasso, Random Forest [RF], Gradient Boosting [GB], Multi-Layer Perceptron [MLP], and their distributional quantile-based extensions, estimating the conditional quantiles.

Model tuning and evaluation. Repeated nested cross-validation[4] (below).



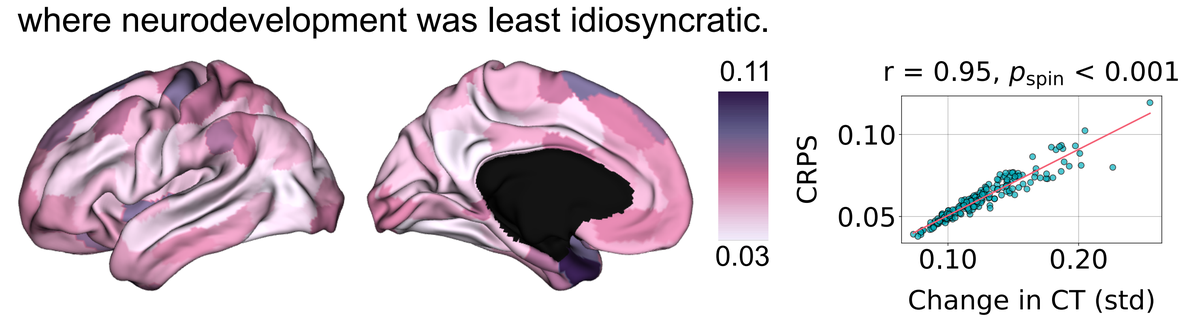
Result

Predictive accuracy. Continuous ranked probability score [CRPS] quantifies prediction error in CT. For mean-regression [point] models, CRPS reduces to mean absolute error [MAE], enabling direct and fair comparison across paradigms[5]. The probabilistic [quantile-based] models outperformed both the null models and the mean-regression [point] models: CRPS improved on average by **56.7%**, **41.4%**, and **31.1%** over Null-i, -ii, and -iii, respectively, whereas mean-regression models failed to significantly improve on Null-iii.

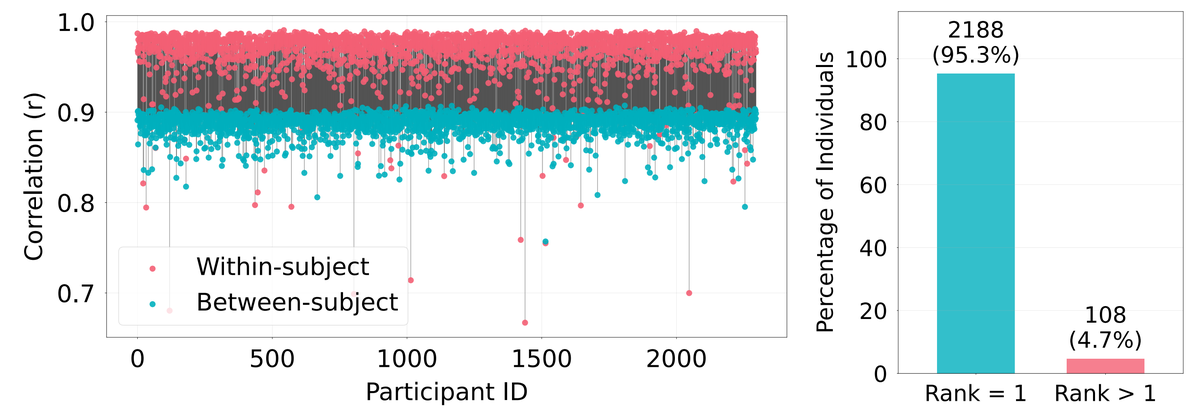


Among the probabilistic [quantile-based] models, QLasso and QRNN generalized most effectively to unseen data, exhibiting the smallest train-test gap, and both produced well-calibrated predictive distributions. Of the two, QRNN won on non-crossing [monotone] quantiles and superior demographic fairness: its residuals and mean probability integral transform [PIT] showed no significant association with any sociodemographic or biological subgroup.

Regional accuracy varied depending on degree of individual variability. Mean CRPS exhibited marked regional heterogeneity and tracked inter-individual variability in CT change: QRNN was most accurate in regions where neurodevelopment was least idiosyncratic.



Individual specificity of predictions. Within-subject correlations were consistently higher than between-subject correlations, indicating that the predicted CT patterns closely mirrored each individual's observed follow-up map [t(2295) = 142.27, p < 0.001]. For **95.3%** of individuals, the correlation between predicted and observed CT across regions was higher for the correct individual than for any other.



Four-year extension. Results are reported for the T01 model and replicated fully for the T02 model, including **55.4%**, **38.4%**, and **29.8%** improvements over Null-i, -ii, and -iii, respectively.

Conclusion

- Our model can enrich baseline assessments by extracting latent signatures of neurodevelopment, as evidenced by superior performance to null models and the individual specificity of predictions.
- Using a probabilistic model allowed predictive uncertainty to capture differences in variability and size of the sociodemographic or biological subgroups, while producing unbiased point-wise estimates.

References. **(1)** Di Biase MA, et al. Mapping human brain charts cross-sectionally and longitudinally. PNAS. 2023 **(2)** Cannon AJ. Non-crossing nonlinear regression quantiles by monotone composite quantile regression neural network, Stoch Environ Res Risk Assess. 2018 **(3)** Tagasovska N, et al. Single-model uncertainties for deep learning. Advances in neural information processing systems. 2019 **(4)** Sasse L, et al. Overview of leakage scenarios in supervised machine learning. J Big Data. 2025 **(5)** Gneiting T, et al. Strictly proper scoring rules, prediction, and estimation. Journal of the American statistical Association. 2007.

Corresponding author. Z.Moradimanesh@FZ-Juelich.de / ZRMOR