Evaluating SynthSeg's Pediatric Brain Segmentations: Longitudinal Volume Assessments, Preprocessing Effects, and Guidelines for Improved Accuracy

Andjela Dimitrijevic[1,2], Alexia Mahlig[1], Fanny Dégeilh[3], and Benjamin De Leener[1,2,4]

[1]NeuroPoly Lab, Institute of Biomedical Engineering, Polytechnique Montréal, Montreal, QC, Canada, [2]Research Center, CHU Sainte-Justine Hospital University Centre, Montreal, QC, Canada, [3]IRISA UMR 6074, EMPENN ERL U-1228, Université de Rennes, CNRS, Inria, Inserm, Rennes, France, [4]Computer Engineering and Software Engineering, Polytechnique Montréal, Montreal, QC, Canada

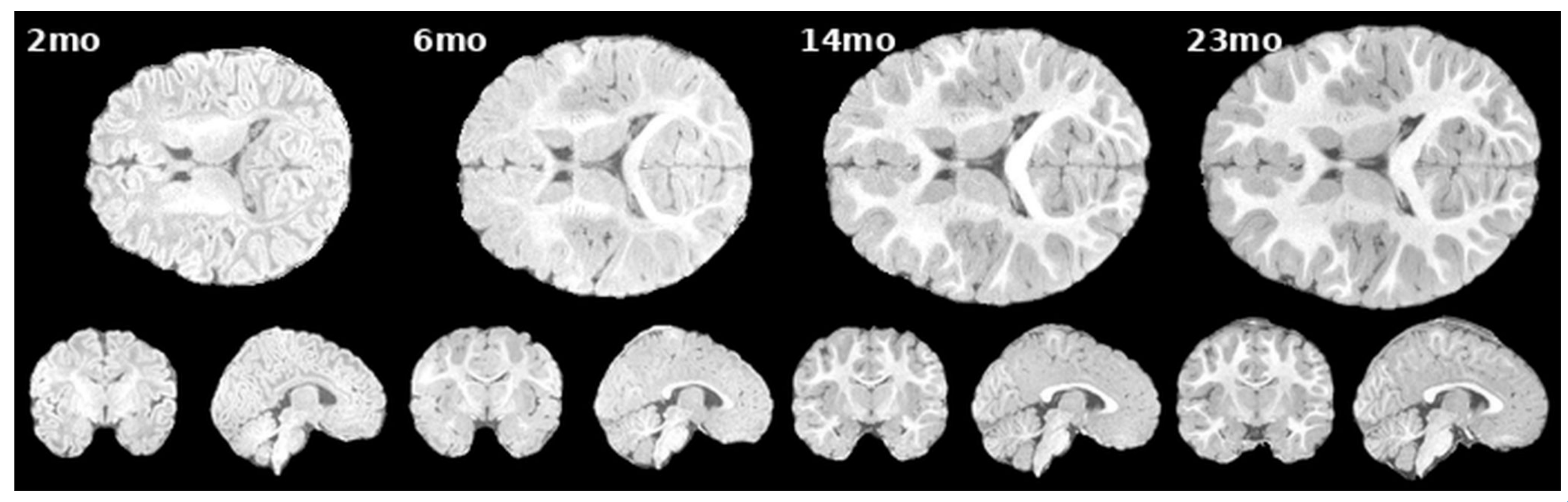
Introduction

★ **Challenge:** Rapid neurodevelopment and frequent motion/contrast artifacts in early-childhood MRI complicate segmentation [1].

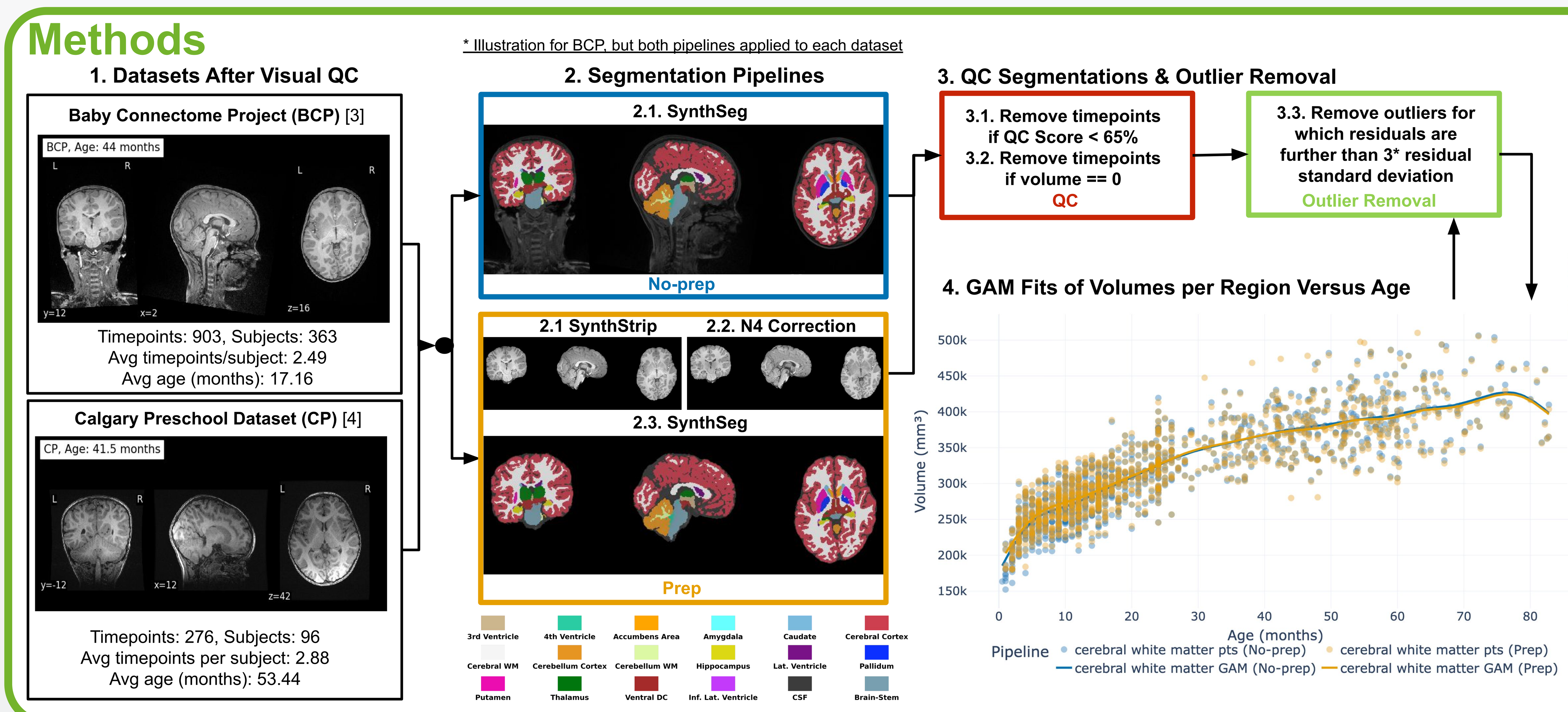
★ **Gap:** SynthSeg (FreeSurfer v7.3.2) [2], an automatic segmentation tool works well in adults, but pediatric segmentation best practices and how MRI pipelines affect longitudinal consistency are poorly defined.

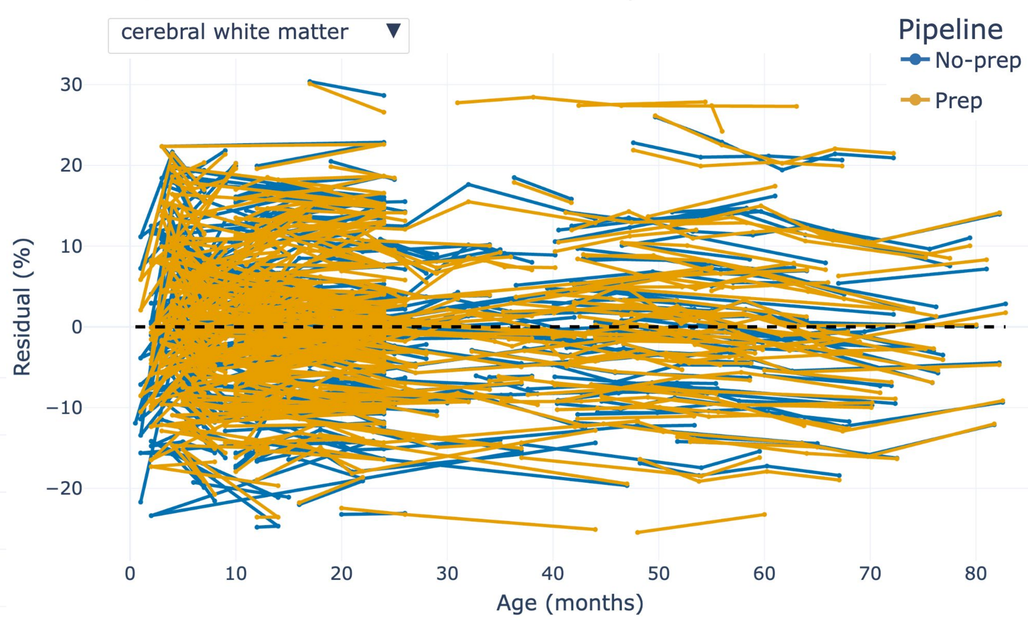
★ **Context:** Gray matter volume rises steeply in infancy (≈1–2 %/month), while white matter myelination follows a more gradual curve, underscoring expected growth trajectories.

★ **Aim:** Benchmark 18 regions in two longitudinal cohorts; test no-preprocessing (**No-prep**) vs N4+skull-strip (**Prep**) pipelines + evaluate how individual trajectories align with established neurodevelopmental patterns.



Methods





★ **RMSD %** is a single-number summary of how much an individual's measured volumes deviate, on average, from the values the GAM predicts for their age. For each subject:

1. **Residual** = Observed – Predicted volume
2. **% Residual** = 100 × Residual / Predicted
3. **RMSD %** = √mean((% Residual)²)

Results & Discussion

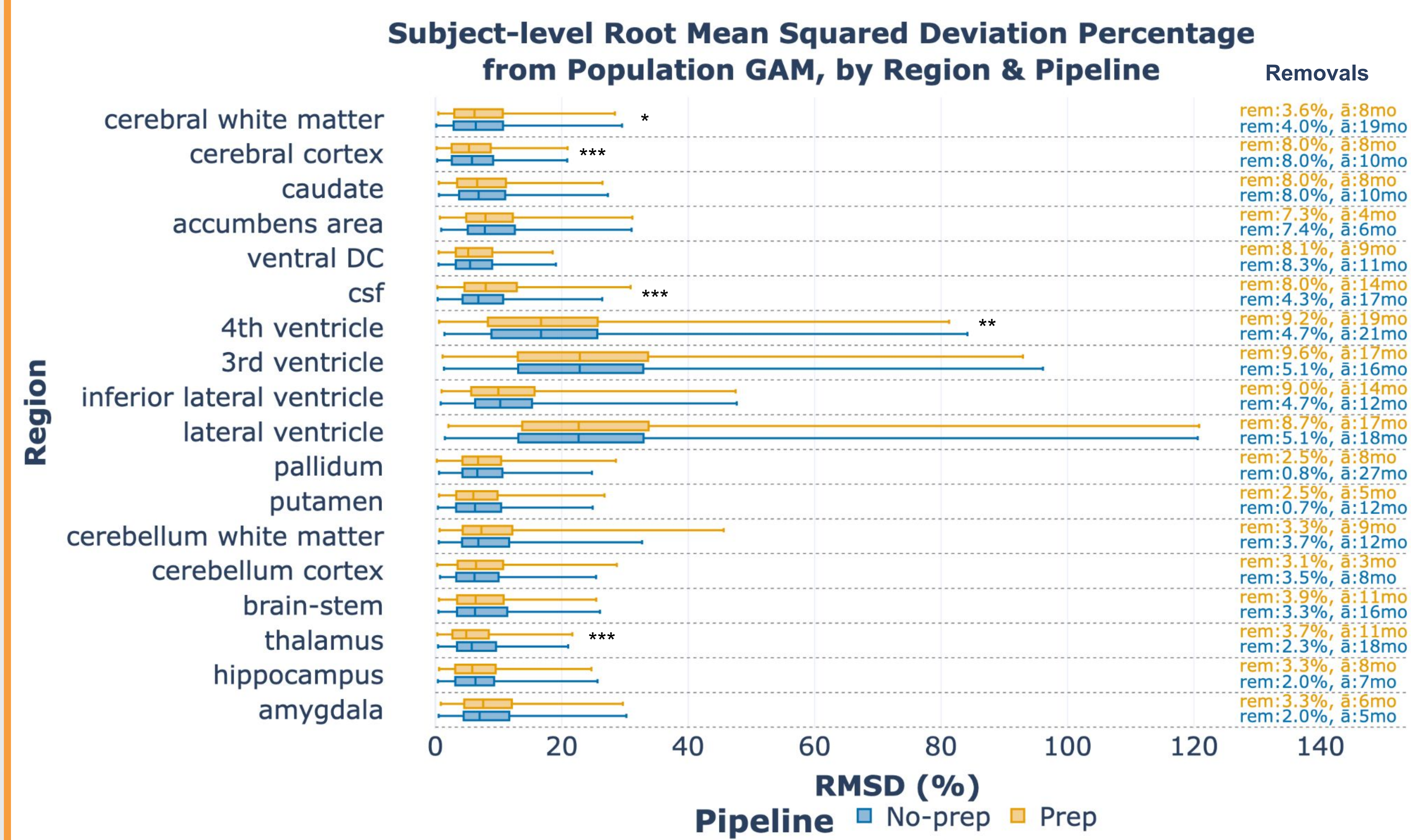


Fig 1. Regional RMSD (%) from the population GAM, compared between No-prep and Prep pipelines. Boxplots show the distribution of % deviation. Adjacent annotations show the removal rate (% of initial sessions) and the mean removal age (months). Significance between pipelines per region is denoted by * ($p < 0.05$), ** ($p < 0.01$) and *** ($p < 0.001$).

★ **Low RMSD %** <=> individual trajectory closely follows population-average neurodevelopmental curve;

★ **High RMSD %** <=> greater departures from the expected developmental pattern.

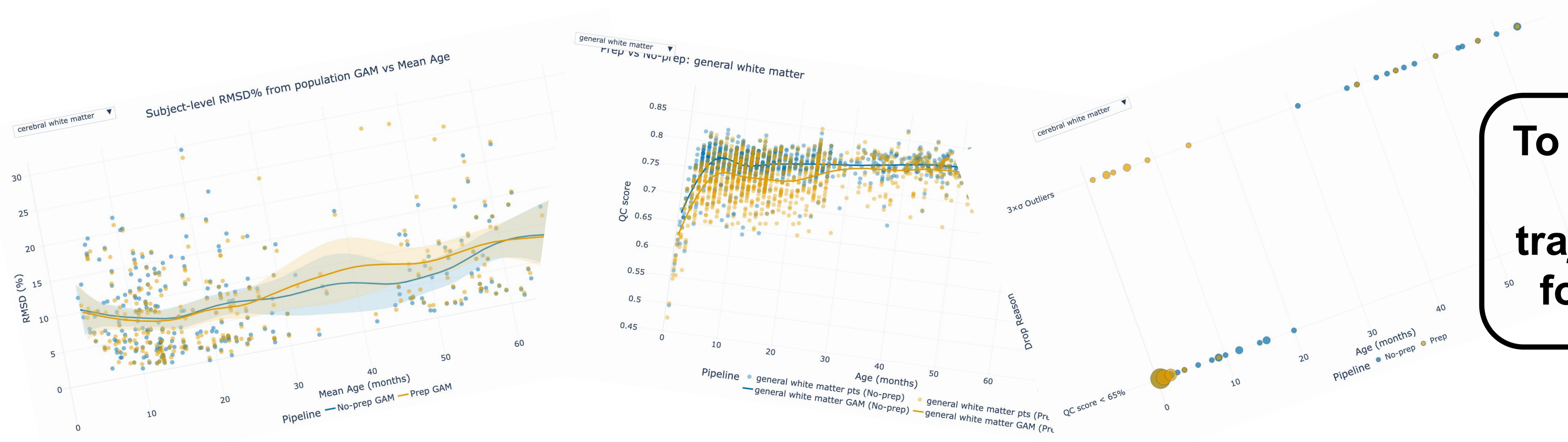
Across 18 regions, the **Prep** pipeline showed lower mean RMSD% in 11 regions, indicating modestly better alignment to the population model.

Largest improvements (Prep vs No-prep):
- ✔ Inf. lat. ventricle: 7.96 % vs 11.74 %
- ✔ Thalamus: 6.14 % vs 6.85 %
- ✔ Putamen: 7.18 % vs 7.41 %
- ✔ Caudate: 7.94 % vs 8.10 %

Regions favoring No-prep (higher RMSD% with **Prep**):
- ✔ Lateral ventricle: 27.29 % vs 26.32 %
- ✔ Cerebellum white matter: 9.10 % vs 8.65 %
- ✔ CSF: 9.26 % vs 8.17 %

Statistical significance

Only 5 regions showed significant Prep vs No-prep differences (Wilcoxon or t-test, $p<0.05$):
- ✔ 4th ventricle, cerebral cortex, cerebral white matter, CSF, thalamus



To explore more interactive figures on the QC removal process, longitudinal trajectories, GAM fits on volumes per age for all 18 regions, scan this QR code →

Conclusion

★ Prep improved alignment (lower mean RMSD %) in 11/18 regions, but performance gains were modest (≤ 1 % difference in most).

★ QC-based removals occurred primarily in very young infants (≈ 1 mo), with more Prep removals in 5/8 broad anatomical groups, especially for general CSF, suggesting SynthStrip's contrast changes hurt its performance in neonates.

★ GAM models of RMSD % versus age show a clear downward trend: residual deviations shrink with increasing age, indicating more stable segmentation in older subjects.

★ **Next steps:** Benchmark SynthSeg against pediatric-focused tools (e.g., iBEAT), include sex as a covariate in developmental models, and develop additional preprocessing or QC strategies tailored for subjects < 10 months, where current methods underperform.

References

[1] Turesky, T. K. et al. *Imaging the Rapidly Developing Brain: Current Challenges for MRI Studies in the First Five Years of Life*. Developmental Cognitive Neuroscience. 2021; [2] Billot, B. et al. *SynthSeg: Segmentation of Brain MRI Scans of Any Contrast and Resolution without Retraining*. Medical Image Analysis. 2023; [3] Howell, Brittany R. et al., *The UNC/UMN Baby Connectome Project (BCP): An Overview of the Study Design and Protocol Development*. NeuroImage. 2019; [4] Reynolds JE et al., *Calgary Preschool magnetic resonance imaging (MRI) dataset*. Data Brief. 2020;

Acknowledgements

This study was supported by Polytechnique Montréal, by the Canada First Research Excellence Fund, by CRSNG, and by the Research Centre of the Sainte-Justine University Hospital.

Contact