Benchmarking Reconstruction Methods for Bundle Segmentation in Single-Shell Diffusion MRI

Amelie Rauland[1,2], Steven L. Meisler[3], Aaron Alexander-Bloch[3,4], Joëlle Bagautdinova[3], Erica B. Baller[3], Raquel E. Gur[3], Ruben C. Gur[3], Audrey C. Luo[3], Tyler M. Moore[3], Oleksandr V. Popovych[1,5], David Roalf[3], Kathrin Reetz[2], Valerie J. Sydnor[6], Simon B. Eickhoff[1,5], Matthew Cieslak[3]*, Theodore D. Satterthwaite[3]*

[1] Research Center Jülich, Jülich, Germany, [2] RWTH Aachen University, Aachen, Germany, [3] University of Pennsylvania, Philadelphia, PA, USA, [4] The Children's Hospital of Philadelphia, Philadelphia, PA, USA, [5] Heinrich Heine University Düsseldorf, Düsseldorf, Germany, [6] University of Pittsburgh, Pittsburgh, PA, USA, * shared senior authors who contributed equally

Contact: a.rauland@fz-juelich.de

Introduction

High-quality research diffusion MRI (dMRI) scans are time- and resource-intensive to acquire.

Legacy dMRI datasets and anonymized hospital-acquired dMRI scans with lower angular resolution offer a valuable, cost-efficient alternative to complement research datasets.

Unclear how reliable metrics extracted from these lower angular resolution scans are.

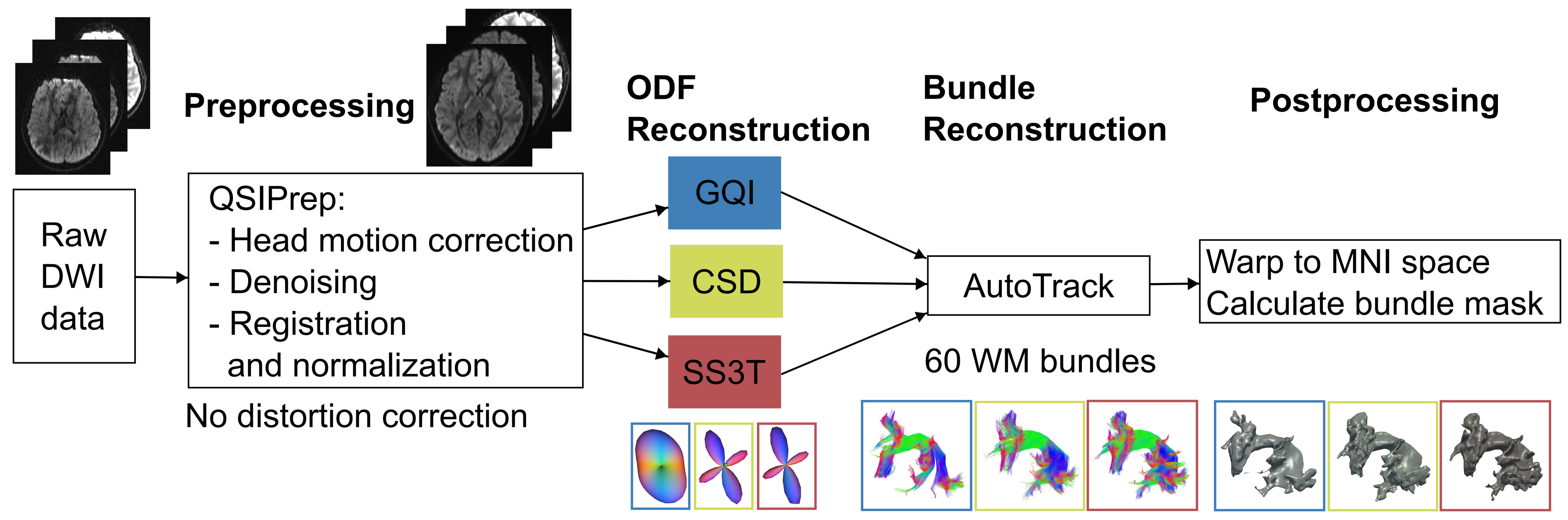
Leverage research dataset with two low angular resolution scans (32 directions) per subject to benchmark reliability of extracted white matter (WM) bundles.

Aim 1: Assess reliability and 'completeness' of reconstructed WM bundles.

Aim 2: Assess ability of features from extracted WM bundles to predict phenotypes, here, cognition.

Aim 3: Compare three methods for ODF reconstruction in single shell low angular resolution data (GQI, CSD and SS3T).

Methods



Data: 1221 subjects from the Philadelphia Neurodevelopmental Cohort with two 32 direction dMRI scans each [1].

Reliability Analysis

- Reconstruction fractions
- Within vs. between subject dice scores
- Discriminability [2]
- Bundle 'completeness': Sensitivity vs. specificity
- Feature ICC

Prediction Analysis

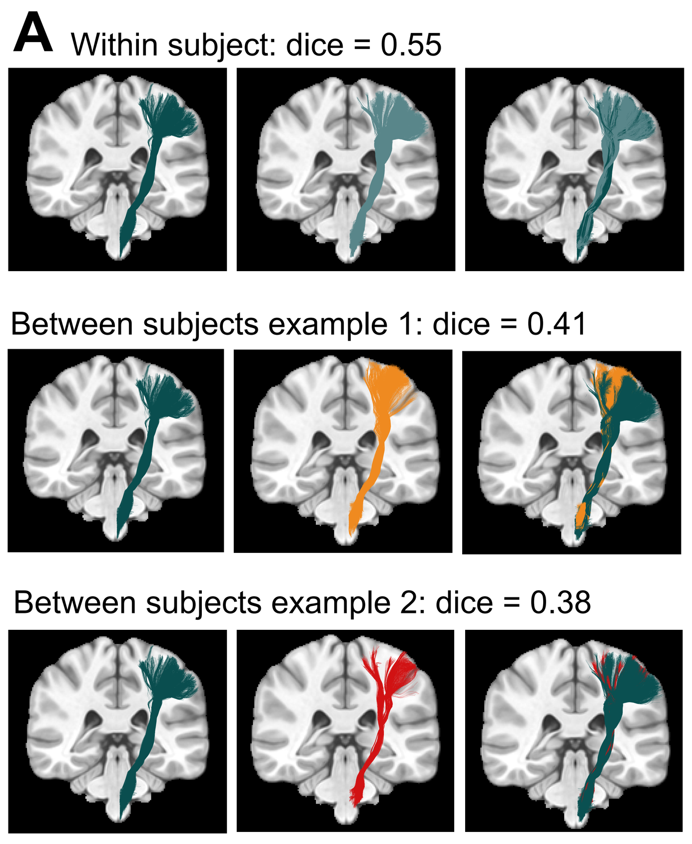
Prediction of complex reasoning using features of the reconstructed bundles (volume, mean MD and mean FA) with a linear ridge regression model.

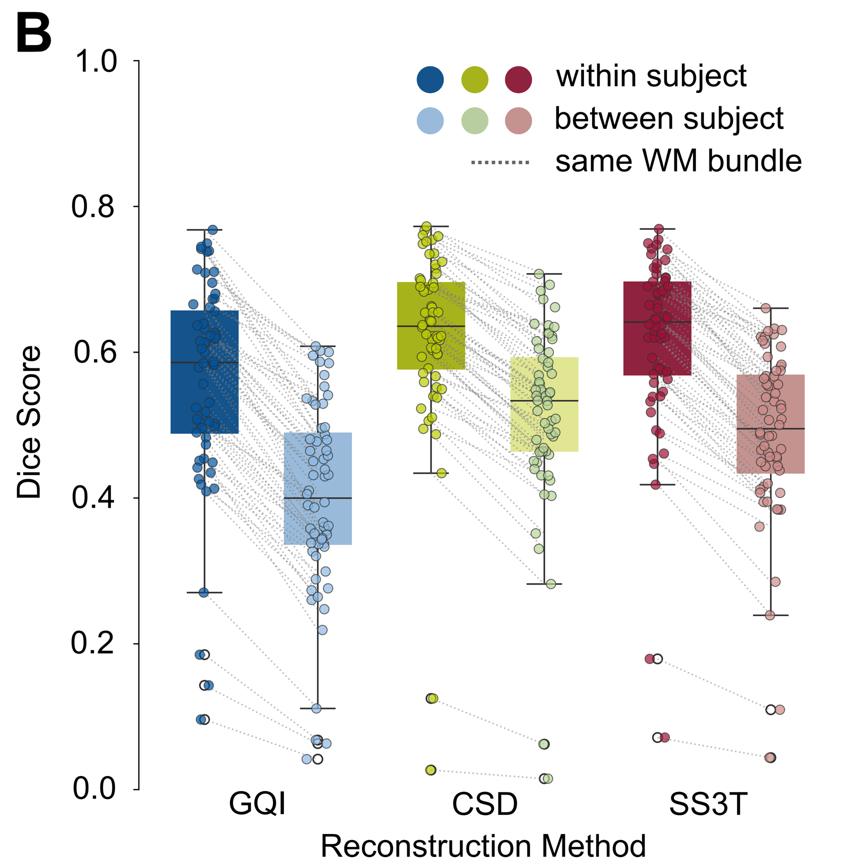
Results

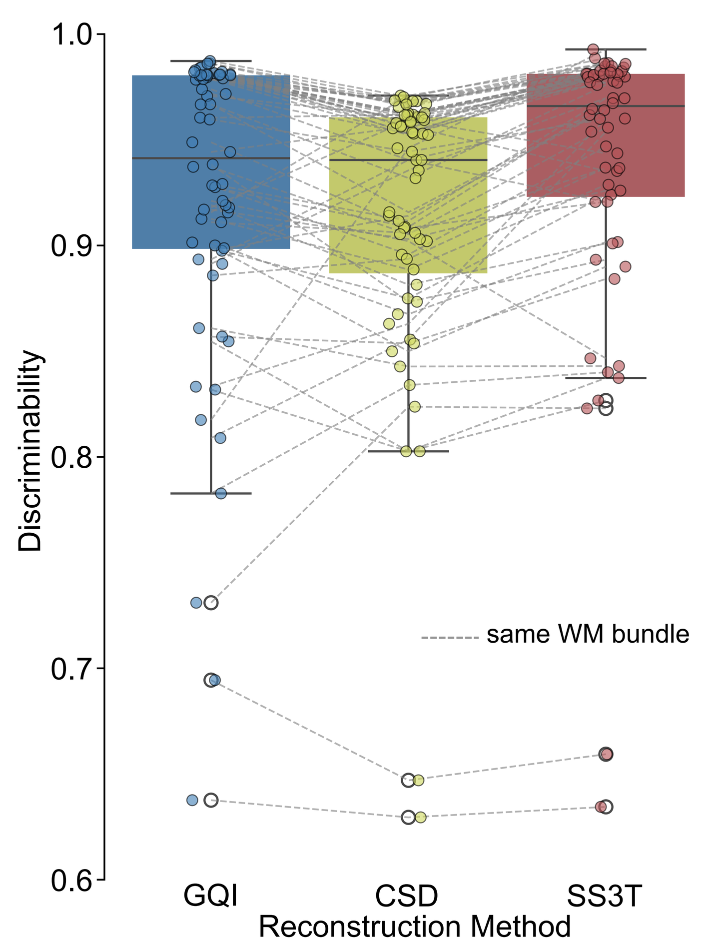
Reliability Analysis

Reconstruction fractions: The majority of bundles was reconstructed with fractions close to 1.

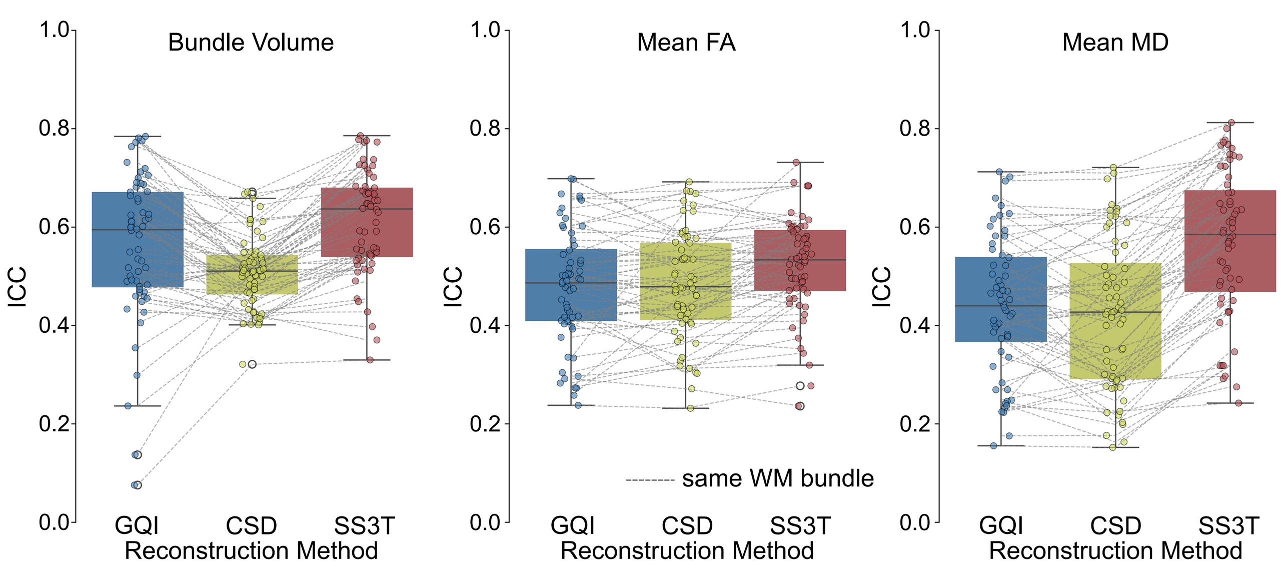
Dice scores: On average across subjects within-subject dice scores > between-subjects dice scores for all WM bundles.







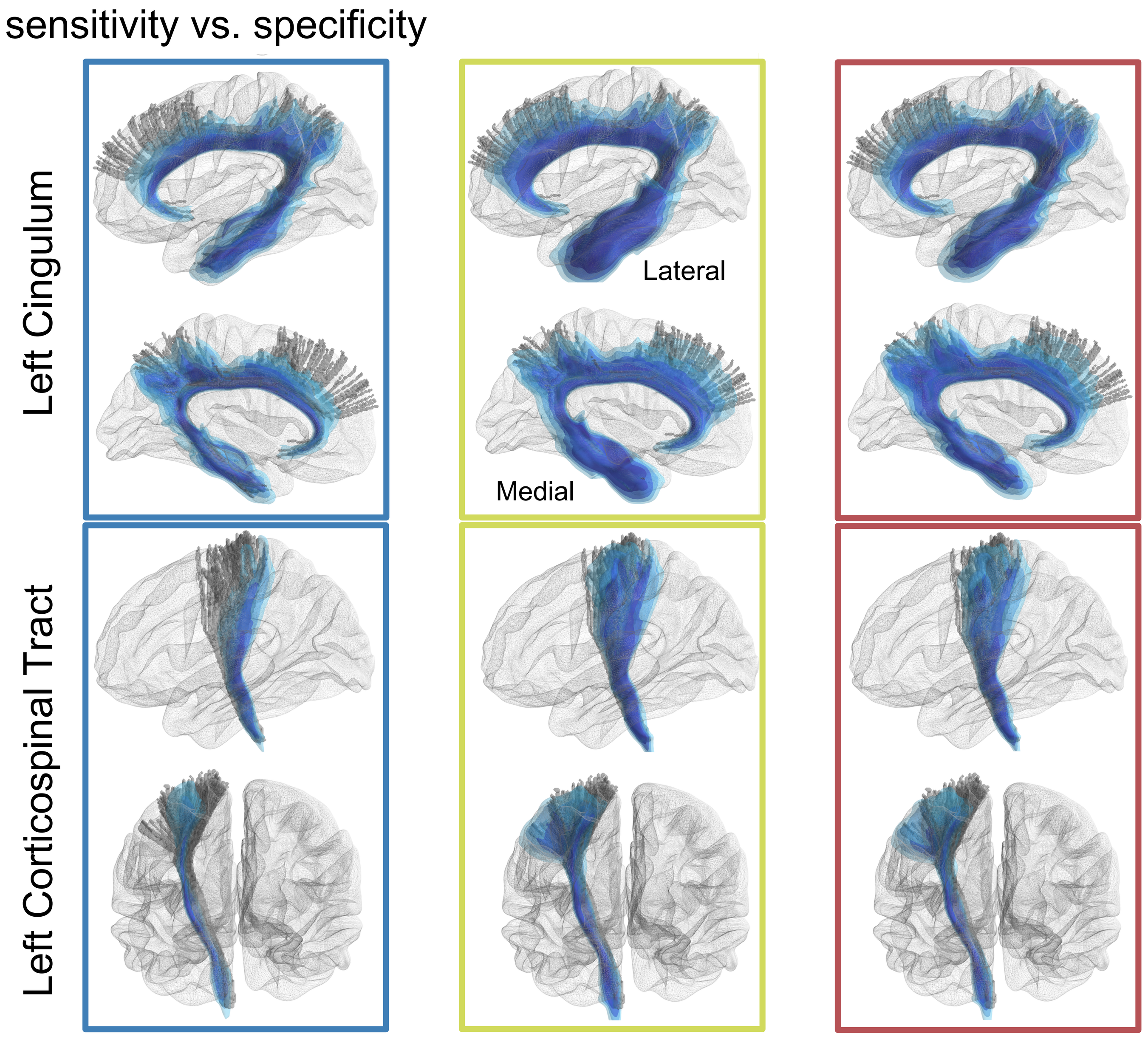
Discriminability: High median discriminability (>0.94) across WM bundles.



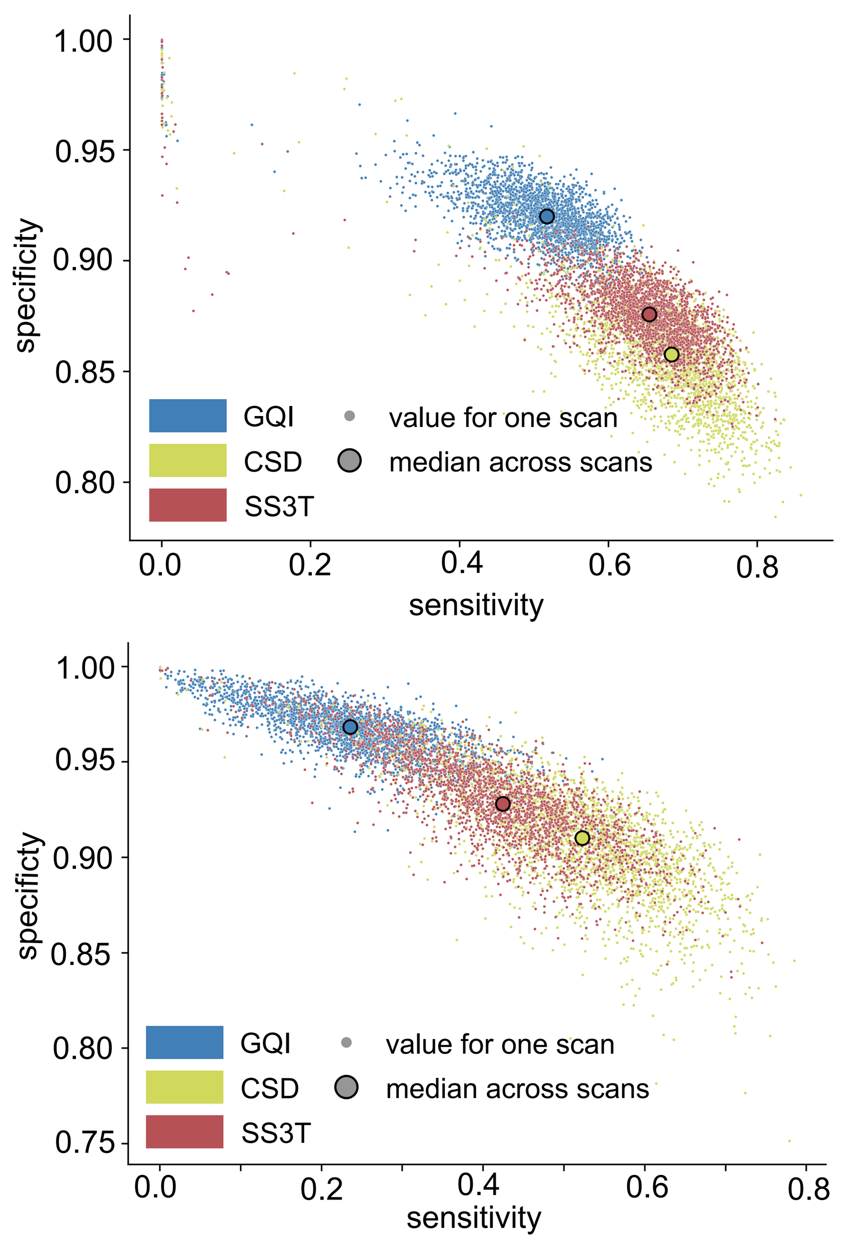
Feature reliability:

- On average, fair reliability of bundle features in terms of intraclass correlation coefficient (ICC).
- ICC is highest for SS3T for all three features.
- Discussion: Slight decrease compared to ICC for WM bundles from high-quality dMRI [5], comparable or better than ICC of the FC [6].

Bundle 'completeness': GQI: sensitivity ↓, specificity ↑, CSD: sensitivity ↑, specificity ↓, SS3T: tradeoff sensitivity vs. specificity

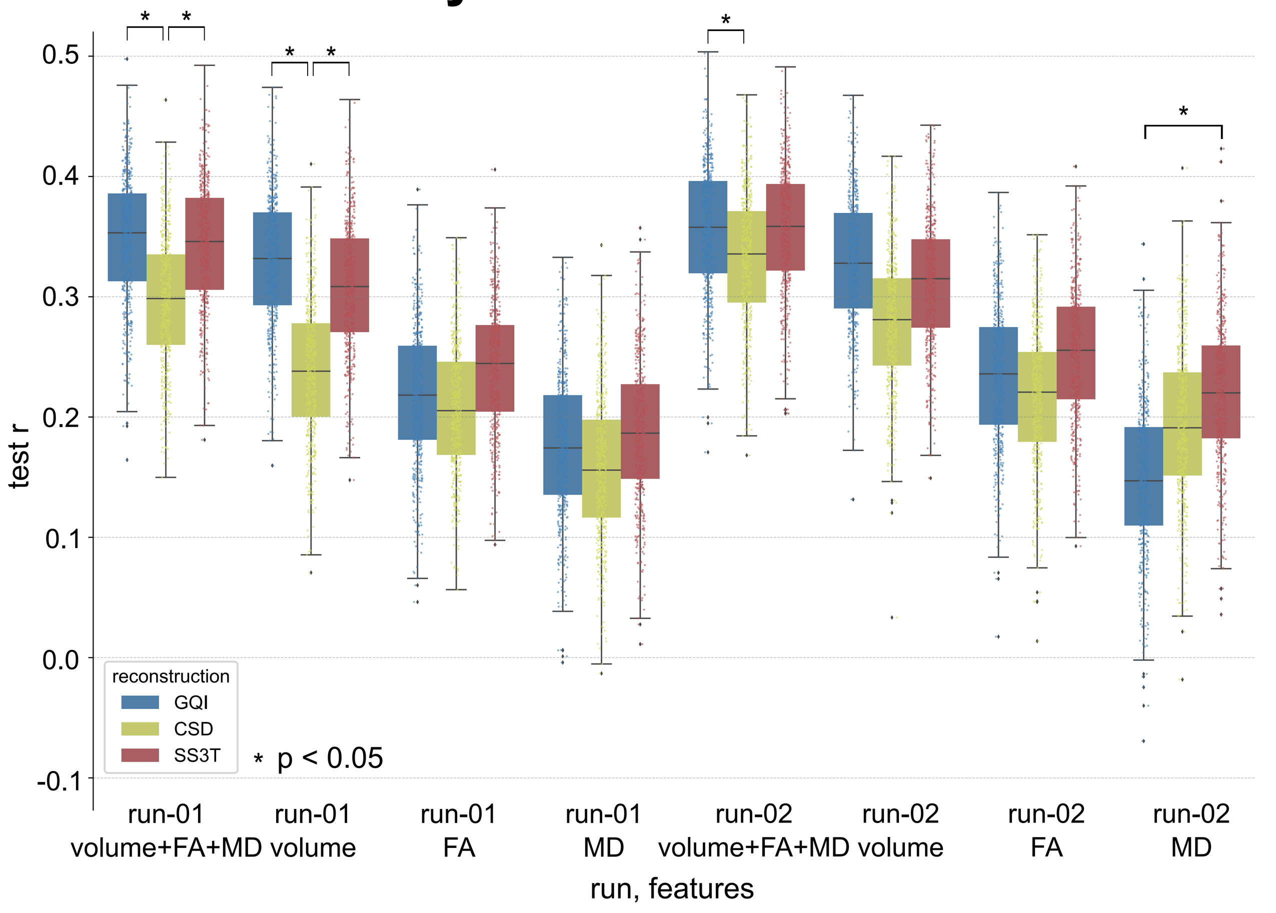


Population map (blues) over atlas bundle (gray).



Sensitivity vs. specificity

Prediction Analysis



Prediction accuracy:

- Stable prediction of complex reasoning (r in 0.15-0.36)
- GQI and SS3T outperform CSD in terms of prediction accuracy.
- All features > volume only > FA only > MD only.
- Discussion: Comparable to prediction accuracy from studies using high-quality dMRI data to predict cognition [3, 4].

Conclusion

- Most WM bundles could be reliably reconstructed from 32-direction, single-shell, clinically feasible dMRI acquisitions.
- SS3T outperformed GQI and CSD, leading to reliable, complete WM bundles suited for predicting brain-behavior relationships.
- Robust prediction of complex reasoning from features extracted from clinically feasible dMRI scans
- Bundle segmentation can achieve robust performance even on lower angular resolution, single-shell dMRI: enormous research potential for dMRI collected in healthcare settings and dMRI legacy datasets.

References: [1] Satterthwaite, Theodore D., et al. "Neuroimaging of the Philadelphia neurodevelopmental cohort." Neuroimage 86 (2014): 544-553. [2] Wang, Zeyi, et al. "Statistical analysis of data repeatability measures." International Statistical Review (2024). [3] Lo, Yui, et al. "The shape of the brain's connections is predictive of cognitive performance: an explainable machine learning study." Human Brain Mapping 46.5 (2025): e70166. [4] Liu, Wan, et al. "Fiber tract shape measures inform prediction of non-imaging phenotypes." arXiv preprint arXiv:2303.09124 (2023). [5] Yeh, Fang-Cheng. "Shape analysis of the human association pathways." Neuroimage 223 (2020): 117329. [6] Noble, Stephanie, et al. "A decade of test-retest reliability of functional connectivity: A systematic review and meta-analysis." Neuroimage 203 (2019): 116157.

Acknowledgments: This work was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) – grant number 269953372/GRK2150.