Title

Structural and functional connectivity reconstruction with CATO - A Connectivity Analysis TOolbox

Short Title

CATO toolbox for structural and functional connectivity reconstruction

Authors

Siemon C. de Lange[1,2] and Martijn P. van den Heuvel[1,3]

Corresponding Author

Siemon C. de Lange, Boelelaan 1085, W&N B-651, 1081 HV, Amsterdam

E-mail address: s.c.de.lange@vu.nl

Affiliations

[1] Department of Complex Trait Genetics, Center for Neurogenomics and Cognitive Research, Vrije Universiteit Amsterdam, Amsterdam Neuroscience, Amsterdam, The Netherlands

[2] Department of Sleep and Cognition, Netherlands Institute for Neuroscience, Amsterdam, The Netherlands

[3] Department of Child Psychiatry, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam Neuroscience, Amsterdam, The Netherlands

Abstract

We describe a Connectivity Analysis TOolbox (CATO) for the reconstruction of structural and functional brain connectivity based on diffusion weighted imaging and resting-state functional MRI data. CATO is an integrative and modular software package that enables researchers to run end-to-end reconstructions from MRI data to structural and functional connectome maps, customize their analysis and utilize different software packages during the data preprocessing. The structural and functional connectome maps can be reconstructed with respect to user-defined (sub)cortical atlases providing aligned connectivity matrices for integrative multimodal analyses. We outline the structural and functional processing pipelines in CATO, the implementation in MATLAB and associated stand-alone application, and the calibration of performance with respect to simulated diffusion weighted imaging data and resting-state functional MRI data from the ICT2015 challenge and the Human Connectome Project. CATO is free open-source software and available at www.dutchconnectomelab.nl/CATO.

Introduction

Network and connectome analysis of macroscale structural and functional brain connectivity data has become a widely used method in neuroscience (Bassett and Sporns 2017). Using neuroimaging data, including magnetic resonance imaging (MRI), one can make a reconstruction of the brain's anatomical connections and functional synchronization patterns (Hagmann, Cammoun et al. 2008) and apply network science to study the topological structure of the derived networks (Rubinov and Sporns 2010). Network studies of the human and animal brain have led to the discovery of general organizational principles of healthy

brain structure and function (Bullmore and Sporns 2012) and abnormalities related to psychiatric and neurological disorders (Fornito, Zalesky et al. 2015).

The availability of multiple software packages fosters the reliability of neuroimaging science. Software packages for the reconstruction and analysis of diffusion weighted imaging (DWI) and resting-state functional MRI (rs-fMRI) derived structural and functional connectivity vary in their objectives, allowing researchers to use software that uses their preferred reconstruction methodology and programming language (Taylor and Saad 2013) (see Table 1 for an overview of software packages). In parallel to open datasets improving the reproducibility in neuroscience by enabling validation of findings across datasets (Mennes, Biswal et al. 2013), variety in methodology permits replication of results using independent analysis pipelines enhancing replicability (National Academies of Sciences and Medicine 2019).

We describe CATO – Connectivity Analysis TOolbox – a flexible toolbox for reconstructing structural and functional connectomes from DWI and rs-fMRI data. CATO is MATLAB based, has a modular structure to allow customized reconstruction of both structural and functional connectomes, and integrates well with respected MRI preprocessing software (Fischl 2012). The toolbox reconstructs structural connectivity using fiber tracking and functional connectivity using Pearson's correlation analysis and integrative multimodal analysis of structural and functional connectivity is facilitated by the option to reconstruct structural and functional connectivity with respect to the same atlas templates. Sharing the software and source code, CATO aims to present an easy, customizable and validated

connectivity toolbox for network neuroscience researchers. The toolbox is completely open-source and free-to-use under the MIT License and is available, with user-manuals and 'how-to descriptions' descriptions, at www.dutchconnectomelab.nl/CATO and GitHub www.github.com/DutchConnectomeLab/CATO.

Methods

CATO is available as a package for MATLAB and as a compiled standalone executable for use on high-performance computer clusters. It has a modular design with separate pipelines for reconstructing structural and functional connectivity. The pipelines consist of multiple processing steps and in what follows, we describe the steps in the structural (part I) and functional (part II) pipelines, followed by a description of the Configuration Assistant that provides a graphical user interface for creating and modifying configuration files for the separate pipelines (see Figure 1). In the Results section, we benchmark and discuss performance with respect to the standard dataset of the ISMRM 2015 Tractography challenge (ITC2015) (Maier-Hein, Neher et al. 2017) and functional data of the Human Connectome Project (Glasser, Sotiropoulos et al. 2013).

Part I: Structural pipeline

Reconstruction parameters for the structural pipeline can be specified on the command line or in a configuration file in JSON-format. The reconstruction parameters describe e.g. the desired processing steps (detailed below), location of the anatomical Freesurfer directory, location of the DWI files, location of the gradient information files, cortical atlas types, and

settings for preprocessing, diffusion reconstruction, fiber tracking and network reconstruction.

Structural preprocessing. The structural pipeline requires pre-processed (high-resolution) whole-brain T1 data (for tissue classification and cortical parcellation) and a (high-resolution) DWI set (we recommend 30 or more diffusion directions (Jones 2004)). First, structural pre-processing of the MRI data is required. We recommend preprocessing of the T1 data by running the steps of the automatic Freesurfer cortical reconstruction process (Fischl, Van Der Kouwe et al. 2004) prior to running CATO, to provide all required files needed for downstream network reconstruction. For the DWI data, we recommend the use of the elaborate FSL package for preprocessing of the DWI data (Jenkinson, Beckmann et al. 2012). Required structural preprocessing steps can be executed by a bash script specified by the parameter `preprocessingScript` allowing users to customize a preprocessing script that fits best to their data, or exclude preprocessing all together and start with already pre-processed data. Three example FSL preprocessing scripts (FSL topup-eddy-preprocessing, FSL eddy-preprocessing and minimal-preprocessing) are provided with the structural pipeline that perform all the needed steps and data output files for subsequent fiber tracking and network reconstruction. These example preprocessing scripts call the well-documented and thoroughly evaluated FSL `topup` and FSL `eddy` tools to correct for susceptibility induced distortions, eddy current distortions and motion artifacts in the DWI data (Andersson and Sotiropoulos 2016), update the b-vectors to adjust for DWI corrections (Leemans and Jones 2009), compute a DWI reference image based on the corrected diffusion-unweighted (b0) volumes, compute the registration matrix between DWI reference image and the anatomical

T1 image using FSL's `bbregister` (Greve and Fischl 2009) and registers the Freesurfer segmentation to the DWI reference image.

Anatomical steps - Parcellation. `Parcellation` creates cortical parcellations of the surface with respect to reference atlases. Atlases standardly applied include the Desikan-Killiany atlas present in Freesurfer (Desikan, Ségonne et al. 2006), combined with the 120, 250 and 500 regions Cammoun sub-parcellations of the Desikan-Killiany atlas (Cammoun, Gigandet et al. 2012) and the Von Economo-Koskinas cortical region and cortical-type atlas (Scholtens, de Reus et al. 2016). Template-specific parcellation scripts create template files in the Freesurfer directory and a `parcellationFile` in the structural pipeline output directory using Freesurfer's `mris_ca_label` to create annotation files, Freesurfer's `mri_aparc2aseg` to merge the cortical parcellation with the subcortical ASEG parcellation and Freesurfer's `mris_anatomical_stats` to provide for every template an anatomical stats file describing for each region, among others, the number of vertices, surface area (mm^2), gray matter volume (mm^3) and average thickness (mm) and Freesurfer's `mri_label2vol` to map the Freesurfer annotation file to the DWI reference image. Easy implementation of other surface based or volume based atlases is possible by adding other atlases to the templates directory.

Anatomical parcellation - Collect region properties. The `collect_region_properties` step collects volumetric and surface data of brain regions and summarizes this data in the `regionPropertiesFile`. The following statistics are included for every brain region: *i)* center of mass of each region (calculated from

the parcellation file), *ii)* the number of vertices, *iii)* surface area (mm^2), *iv)* gray matter volume (mm^3), and *v)* average thickness (mm) (all from Freesurfer's stats file). These metrics are used in later steps of the pipeline.

Reconstruction diffusion signal. The `reconstruction_diffusion` step estimates the white matter fiber organization in each voxel from the measured DWI data. Structural connectivity modelling is based on the principle that white matter fibers restrict the movement of water molecules resulting in peaks (preferred directions) in the diffusion patterns of water molecules in voxels. The `reconstruction_diffusion` step provides three methods to infer the diffusion peaks from DWI data, including Diffusion Tensor Imaging (DTI), Constrained Spherical Deconvolution (CSD) and Generalized Q-sampling (GQI) and the user can easily implement other methods when required. For each method, the reconstructed peaks of all voxels are saved in the `diffusionPeaksFile` and associated diffusion measures are saved in the `diffusionMeasuresFile`.

Diffusion Tensor Imaging (DTI). `DTI` reconstruction models the signal measured in a voxel by a single tensor that defines the diffusion profile associated with one preferred diffusion-direction. Tensor estimation is performed using the informed RESTORE algorithm (Chang, Jones et al. 2005) that performs tensor estimation while identifying and removing outliers during the fitting procedure, reducing the impact of physiological noise artifacts on the diffusion tensor modelling. The resulting set of diffusion gradients corresponding to non-outlying measurements is evaluated for being well-balanced by automatically selected thresholds obtained using bootstrapping (de Reus 2015). The Levenberg-Marquardt method (based on implementation by Gavin (Gavin 2019) is used to solve the nonlinear least squares

problem. Based on the reconstructed tensor, diffusion measures fractional anisotropy, axial diffusivity, radial diffusivity and mean diffusivity are computed (Alexander, Lee et al. 2007).

Constrained Spherical Deconvolution (CSD). `CSD` reconstructs the diffusion peaks in a voxel by deconvolution of the measured signal with the diffusion profile associated with a fiber (Tournier, Calamante et al. 2007). Signal deconvolution is performed in the super-resolution spherical harmonics framework to allow for a natural description of the surface of a diffusion process. The spherical deconvolution is constrained to eliminate spherical harmonics with negative values which reduces high frequency noise and results in the reconstruction of a few well-defined peaks (Tournier, Calamante et al. 2007). The implementation is based on the CSD reconstruction method in the Dipy software package (Garyfallidis, Brett et al. 2014) and as described in Tournier et al. (2007).

Generalized Q-sampling Imaging (GQI). `GQI` reconstructs the diffusion peaks in a voxel by utilizing the Fourier transform relation between the diffusion signal and the underlying diffusion displacement (Yeh, Wedeen et al. 2010). The implementation follows the method described in Yeh et al. (Yeh, Wedeen et al. 2010) and the example code as presented on http://dsi-studio.labsolver.org. Diffusion measures generalized fractional anisotropy (Tuch 2004) and quantitative anisotropy (Yeh, Wedeen et al. 2010) of each peak are also calculated and added to the `diffusionMeasuresFile`.

Combined implementation. Comparing the three reconstruction methods, DTI is a robust and relatively simple method, with CSD and GQI being more advanced diffusion models that allow a better description of the diffusion signal in voxels having a more complex underlying white matter organization. To combine the strengths of simple and advanced reconstruction methods, CATO gives users the option to combine DTI with CSD or GQI. This option

performs DTI modelling (being more robust) in voxels where CSD or GQI estimate only one peak, and uses CSD or GQI in voxels where multiple peaks are detected (being more flexible in voxels with a complex fiber architecture). Specific parameter settings for each reconstruction method are set in the configuration file and are described in more detail in the online documentation.

Intermediate output files. The `reconstruction_diffusion` step provides the user the additional option to export diffusion measures to a NIFTI volume file `diffusionMeasuresFileNifti`. Standard diffusion measures that are exported include fractional anisotropy, axial diffusivity, radial diffusivity and mean diffusivity.

Reconstruction fibers. The `reconstruction_fibers` step performs fiber tracking based on the diffusion peaks in each voxel. The fiber tracking step utilizes the standard "Fiber Assignment by Continuous Tracking" (FACT) algorithm (Mori, Crain et al. 1999) (adjusted for multiple peaks when CSD, GQI or combined methods are used). FACT describes a deterministic tracking algorithm that starts fiber reconstruction from seeds in the white matter and propagates streamlines in the main diffusion axis of the voxel while updating the propagation direction each time the tip of the streamline enters a new voxel. The CATO implementation of FACT starts from one or multiple seeds (number of seeds set by `NumberOfSeedsPerVoxel`) in all voxels with a segmentation label matching the list provided by `startRegions`. Fiber reconstruction stops if a tracker is *i)* in a region with low fractional anisotropy (FA < `minFA`, *ii)* in a stopping region with segmentation code included in the `stopRegions` variable, *iii)* about to revisit the current voxel or previously

visited voxel, *iv)* about to enter a forbidden region with segmentation code included in the `forbiddenRegions` variable or *v)* about to make a sharp turn (with angle > `maximumAngleDeg`).

Reconstruction fiber properties. The `reconstruction_fiber_properties` step identifies fiber segments that connect brain regions and calculates fiber measures as preparation for the network reconstruction step. For each atlas and reconstruction method, the `reconstruction_fiber_properties` step iterates through all fibers and if a fiber crosses two or more brain regions of interest, as defined by the regions of interest file (`ROIsFile`), then the shortest fiber segment between each region pair is included in the fiber properties file (`fiberPropertiesFile`). In addition to the start and end point of each fiber segment and the associated region pair, additional measures for fiber segments are stored, including: *maximum turn angle, minimum fractional anisotropy, fiber length* (physical length in mm) and the average *fractional anisotropy*, *axial diffusivity*, *radial diffusivity*, *mean diffusivity* and *generalized fractional anisotropy.* Diffusion measures are averaged over voxels weighted by the length of the traversed path through each voxel (Tuch 2004).

Reconstruction structural network. The `reconstruction_network` step builds the connectivity matrices for each of the selected (sub)cortical atlases and reconstruction methods. Brain regions included in the connectivity matrix, and their order, are defined by the regions of interest file (`ROIsFile`). The network matrix is constructed by iterating through all fiber segments in the `fiberPropertiesFile` and fiber segments connecting regions of interest are added to the connectivity matrix.

In addition to the fiber reconstruction criteria used in the `reconstruction_fibers` step, fibers can additionally be filtered in the `reconstruction_network` step by their projection length (with only fibers longer than `minLengthMM` being included in the network reconstruction), minimum fractional anisotropy (including only fibers that touch voxels with fractional anisotropy higher than `minFA`) and maximum angle (including only fibers that make turns in their trajectories smaller than `maxAngleDeg`). Connectivity matrices are saved to the file `connectivityMatrixFile`. Connectivity matrices include weights by the *number of streamlines* (i.e. the number of fibers) that connect two regions, *fiber length* (physical length in mm averaged across fibers), the average *fractional anisotropy*, *axial diffusivity*, *radial diffusivity*, *mean diffusivity*, *generalized fractional anisotropy, streamline volume density* and *streamline surface density* (Tuch 2004) of the reconstructed streamline segments.

Part II: Functional connectivity

The functional connectivity pipeline builds upon the same modular organization as the structural connectivity pipeline, including a preprocessing step, two anatomical steps and a connectivity reconstruction step (see Figure 1). Resting-state fMRI processing starts with the main function `functional_pipeline` that runs the processing steps according to the parameters provided on the command line or rs-fMRI pipeline configuration file. The functional pipeline includes the following steps:

Preprocessing. The step `functional_preprocessing` provides users the option to preprocess rs-fMRI data with their preferred methods, specified in a bash script

(`preprocessingScript`). CATO provides an example preprocessing script that includes steps for correcting slice timing using FSL tool `slicetimer` and correcting motion artifacts using FSL tool `MCFLIRT` (Jenkinson, Bannister et al. 2002), computing a rs-fMRI reference image by averaging all (motion corrected) rs-fMRI frames (using FSL tools), computing a registration matrix between the rs-fMRI reference image and the T1 image (Greve and Fischl 2009) (using Freesurfer) and registration of the T1 parcellation to the rs-fMRI image (using Freesurfer). The modular pipeline allows preprocessing with other packages and completely custom preprocessing when required.

Anatomical parcellation. The first steps of the functional pipeline include the same parcellation steps `parcellation` and `collect_region_properties` as the structural pipeline to parcellate the cortical surface with respect to the same atlases as the structural pipeline, such that structural-functional relationships can be explored in post-processing integrative analyses. The anatomical steps further compute the anatomical statistics of these regions and collect the region properties into the region properties file `regionPropertiesFile`.

Computation motion metrics. `Compute_motion_metrics` computes for each frame the motion metric framewise displacement (FD) and the change in signal intensity between frames (known as 'DVARS') (Power, Barnes et al. 2012). Measures are derived from a `motionParametersFile` file that can be created in the preprocessing step with MCFLIRT (Jenkinson, Bannister et al. 2002) and are saved in the `motionMetricsFile`. Framewise displacement (FD) is computed as the sum of the estimated translational and

rotational displacement in a frame (Power, Barnes et al. 2012), with rotational displacement defined in degrees in the `motionParametersFile` file. FD is converted to millimeters by calculating the expected displacement on the surface of a sphere of radius 50 mm as model for the cerebral cortex (Power, Barnes et al. 2012). DVARS is calculated as the square root of the average squared intensity differences of all brain voxels between two consecutive rs-fMRI frames (Power, Barnes et al. 2012).

Time-series processing. The `reconstruction_functional_network` step computes the functional connectivity matrices preceded by (optional) covariate regression, band-pass filtering and scrubbing of the time-series.

Regression. Per voxel, covariates (referred to as regressors) are removed from the rs-fMRI time series by calculating the residuals of a linear model of the signal intensities with the regressors as predictors. Standard regressors include the linear trends of motion parameters, first order drifts of motion parameters and mean signal intensity of voxels in white matter and CSF (specific regions are defined by configuration parameter `regressionMask` and each region code in the regression mask is included as separate regressor). Mean signal intensity of all voxels in the brain can optionally be included as an additional regressor to perform global mean correction (indicated by `globalMeanRegression`) (Fox, Zhang et al. 2009).

Band-pass filtering. Next, (optional, but recommended) band-pass filtering to the rs-fMRI data is applied (indicated by parameter `filter`). A zero-phase Butterworth bandpass filter is used with low-pass and high-pass cutoff frequencies set by the configuration parameter `frequencies`.

Scrubbing. Frames that display significant motion artifacts are removed from the rs-fMRI time-series (scrubbing) (Power, Barnes et al. 2012). When scrubbing is applied (optional, but highly recommended), frames with motion artifacts are identified based on two indicators: *i)* having framewise displacement FD larger than `maxFD` and *ii)* having a DVARS larger than Q3 + `maxDVARS` × IQR, where IQR refers to the interquartile range IQR = Q3 – Q1, with Q1 and Q3 referring to the first and third quartile of the DVARS of all frames. Frames with a number of indicators larger or equal to `minViolations` are labeled as frames with potential motion artifacts and are excluded from further analysis. To accommodate temporal smoothing of data, frames consecutive to frames with labeled motion artifacts are optionally excluded: configuration parameter `backwardNeighbors` determines the number of preceding frames and `forwardNeighbors` determines the number of succeeding frames to be excluded from further analysis.

Network reconstruction. The `reconstruction_functional_network` step computes region-to-region functional connectivity by means of correlation analysis. Functional connectivity is computed between brain regions (specified by the `ROIsFile`) as the Pearson's correlation coefficient of the average signal intensity of these regions across the selected frames. Connectivity matrices are saved to the file `connectivityMatrixFile` and time series of each region are saved to the file `timeSeriesFile`. Multiple `reconstruction_functional_network` steps can be performed to explore different reconstruction settings and are distinguished by using the parameter `methodDescription`.

Configuration assistant

Contributing to the usability of CATO is the Configuration assistant, which is a user-friendly online graphical user interface to create and modify a configuration file with the settings and parameters for the structural and functional pipeline (www.dutchconnectomelab.nl/CATO). Besides easy configuration through visual dialogs, the Configuration assistant also checks all configuration parameters for their validity to reduce errors due to typos and incorrect file names.

Structural pipeline benchmarking

Data. The variety of implemented reconstruction methods enables researchers to select a preferred trade-off between sensitivity and specificity in connectome reconstructions (Zalesky, Fornito et al. 2016). To benchmark CATO to other reconstruction strategies and to evaluate the impact of different diffusion modelling methods and parameters, we compared the used diffusion modelling and fiber reconstruction methodology with data of the ISMRM 2015 Tractography challenge (ITC2015) (Maier-Hein, Neher et al. 2017). ITC2015 allows tractography pipelines to be evaluated with respect to a defined 'ground truth' white matter skeleton of 25 distinct white matter bundles distilled from DWI data of a single Human Connectome Project (HCP) subject by an expert radiologist (Glasser, Sotiropoulos et al. 2013), including cortico-cortical, cortico-spinal and cortico-cerebellar tracts (see Table 2 for a description of the included white matter bundles and Figure 2). This predefined ITC2015 white matter skeleton was used to subsequently simulate DWI images, which were distributed across research groups for analysis. Reconstructed fiber clouds from each research group were collected in the ITC2015 competition for benchmarking. Using CATO (version 3.0), we reconstructed four fiber clouds (for the method DTI, CSD, GQI and GQI-DTI

combined) using default HCP reconstruction parameters and a series of fiber clouds in which reconstruction parameters were adjusted (see Supplementary Methods 1). The reconstructions were benchmarked on two key reconstruction features, including 1) the correctness of the spatial trajectory of reconstructed white matter bundles, and 2) the specificity and sensitivity of the reconstructed connectome networks.

Spatial reconstruction measures. We evaluated the spatial trajectory of the reconstructed fiber clouds with respect to three quality metrics using scripts provided by ITC2015 (Maier-Hein, Neher et al. 2017). The first two metrics are based on clustering fibers into white matter bundles and subsequently computation of their spatial overlap with the ITC2015 ground truth white matter bundles. This provides 1) the number of true-positive reconstructed white matter bundles (i.e. number of valid bundles, VB), 2) the number of reconstructed false-positive bundles (i.e. number of invalid bundles, IB). 3) The third metric counts the number of fibers with trajectories that spatially overlapped with the ground truth white matter skeleton (i.e. number of valid fibers, VF). The ITC2015 referred to this third measure as 'valid connections'; we here refer to it as 'valid fibers' to avoid confusion with connectome reconstruction measures.

Connectome reconstruction measures. We also evaluated the quality of the connectome reconstructions (the core aim of CATO) using the ITC2015 data. To this end, fiber clouds from all ITC2015 submissions were downloaded from the data archive on Zenodo (https://doi.org/10.5281/zenodo.840086) and converted to 87 fiber cloud files in TRK-format using Dipy resulting (9 submissions were excluded due to errors including 15_0, 7_1, 7_2,

7_3, 3_3, 6_0, 6_1, 6_2 and 6_4). The anatomical T1-weighted image of the ITC2015 was parcellated into 68 distinct cortical regions according to the Freesurfer Desikan-Killiany atlas (describing 34 left-hemispheric and 34 right-hemispheric cortical regions, (Cammoun, Gigandet et al. 2012), using Freesurfer version 6.0.0 (Fischl, Van Der Kouwe et al. 2004)). Using this parcellation, we computed 1) a ground truth connectome based on the ITC2015 ground truth white matter tracts, 2) connectome maps based on the fiber clouds from all ITC2015 submissions and 3) connectome maps based on the fiber clouds obtained from CATO reconstructions of the ITC2015 dataset (see Supplemental Methods 1 for details on the CATO reconstruction of the ITC2015 data). All three series of connectome maps were formed by combining the cortical parcellation with the respective fiber clouds into connectivity matrices using the `reconstruction_fiber_properties` and `reconstruction_network` functions, thus using identical processing steps for all connectome reconstructions. Connectome reconstruction quality was then assessed by means of computing the false positive rate (100% × number of false positive connections / number of absent connections in ground truth connectome) and false negative rate (100% × number of false negative connections / number of connections in ground truth connectome) (Zalesky, Fornito et al. 2016) as measured with respect to the ITC2015 ground truth connectome.

Functional pipeline benchmarking

Functional connectomes were benchmarked by assessing test-retest reliability on HCP data of 500 subjects for repeated rs-fMRI scans in 2 sessions (Glasser, Sotiropoulos et al. 2013). We constructed for all subjects and for all scans a connectivity matrix using default HCP reconstruction parameters (see Supplementary Methods 2). The intersession '1-1' intra-class

correlation (ICC) was calculated for each connection across sessions (using the 'LR' scan of the first and second session) as

$$ICC = (MS_b - MS_w) / (MS_b + (n-1)MS_w)$$

to quantify the degree of absolute agreement among scans across subjects, with MS_b the between-subject mean squared strength of the connection, MS_w the within-subject mean squared strength of the connection and n the number of scans ($n = 2$) per subject (McGraw and Wong 1996). To additionally filter out potential spurious connections, ICC scores were also computed of functional connections that were selected by applying a group connectivity mask across subjects and sessions. Connections were selected by a proportional threshold that selected 20% strongest correlations in a group-averaged functional connectome obtained by averaging connectivity matrices across subjects and sessions. ICC scores of functional connectivity were also computed between all region pairs (i.e. on the full connectivity matrix, not applying group mask filtering). We evaluated the ICC of connections for four common cortical atlases included in CATO (Desikan-Killiany atlas, the 120 and 250 subdivisions of the Desikan-Killiany atlas and the Von Economo-Koskinas atlas). The effect of global mean correction, bandpass filtering and scrubbing were also explored by evaluating ICC scores associated with the application of these reconstruction options. Reliability was interpreted using the description by Cicchetti (Cicchetti 1994) stating that ICC < 0.4 reflects poor reliability, ICC = [0.40 - 0.59] reflects fair reliability, ICC = [0.60 - 0.74] reflects good reliability and ICC = [0.75 - 1.00] reflects excellent reliability.

Results

Structural pipeline benchmarking

Spatial reconstruction. Performance was validated in comparison with tractography-based fiber clouds from (in total) 96 submissions of 20 different sites of the ITC2015 challenge. ITC2015 submissions showed a wide range of scores on the spatial reconstruction quality measures: the number of validly reconstructed bundles ranged from 5 to 24 bundles out of the 25 ground-truth bundles and the number of invalid bundles ranged from 27 to 386 bundles (Figure 3a). 22 of the 25 ground-truth bundles (88%) were reconstructed, which was one bundle less than the median of 23 of the ITC2015 submissions (Figure 3a). White matter bundles that were missed included the left cortico-spinal tract, anterior commissure and posterior commissure. Across reconstruction methods, 64 to 78 invalid bundles were traced, which were not included in the ground-truth white matter skeleton (CSD: 78, DTI: 71, GQI-DTI: 78 and GQI: 64), numbers that are comparable to ITC2015 submissions which had a median value of 77 invalid bundles (Figure 3a). The percentage of fibers with tractographies that spatially overlapped with the ground truth white matter skeleton ranged between 3.75% and 92.5% for the ITC2015 submissions (median: 62.6%), with CATO values ranging between 43% and 55% (CSD: 43%, DTI: 55%, GQI-DTI: 56% and GQI: 48%, Figure 3b).

Connectome mapping. We next examined structural connectome reconstruction performance. ITC2015 submissions (67 submissions included) showed a wide range of results, with false positive rates between 2.0% - 92.7% (median = 14.5%) and false negative rates between 0.0% - 55.9% (median = 20.3%). This emphasizes a wide spectrum of specificity-sensitivity trade-offs across the different strategies used by groups participating in ITC2015. CATO connectome reconstructions included false positive rates between 6.7% - 14.9% (CSD: 14.9%, DTI: 8.1%, GQI-DTI: 9.4%, GQI: 6.7%) and false negative rates between 27.0% -

35.9% (CSD: 27.0%, DTI: 30.8%, GQI-DTI: 27.6%, GQI: 35.9%, see Figure 3c). This relative low false positive rate places CATO reconstructions at default settings as a connectome reconstruction tool with a relative conservative reconstruction sensitivity and relatively high connectivity reconstruction specificity (Zalesky, Fornito et al. 2016). Levels of sensitivity and specificity of connectome reconstruction can further be tuned by varying the used reconstruction parameters, with diffusion reconstruction parameters showing small impact and fiber reconstruction parameters having large impact on the sensitivity-specificity trade-off (see Supplementary Methods 3).

Functional pipeline benchmarking

Reproducibility of reconstructed functional connectivity was estimated in a test-retest design based on repeated rs-fMRI scans from the HCP dataset. Reliability of connections in the group connectivity matrix was assessed by ICC, presented for various cortical parcellations in Figure 4a. In the standard Desikan-Killiany atlas, 30.7% of the connections showed 'poor' reliability, 68.4% showed 'fair' reliability and 0.9% 'good' to excellent reliability according to Cicchetti's interpretation (median ICC=0.43, IQR=0.39-0.48). In the 120 regions sub-parcellation of the Desikan-Killiany atlas, 23.1% of the connections showed 'poor', 73.5% 'fair' and 3.3% 'good' to 'excellent' reliability (median ICC=0.44, IQR=0.40-0.49). In the 250 regions sub-parcellation of the Desikan-Killiany, 19.5% of the connections showed 'poor', 72.6% 'fair' and 7.8% 'good' to 'excellent' reliability (median ICC=0.46, IQR=0.41-0.51). In the Von Economo-Koskinas atlas, 28.3% of the connections showed 'poor', 70.9% 'fair' and 0.8% 'good' to 'excellent' reliability (median ICC=0.43, IQR=0.39-0.47). The reliability of functional connectome reconstructions was further examined for various

network reconstruction parameters. Focusing on the Desikan-Killiany 114 regions sub-parcellation, we observed that network reconstructions that included global mean regression resulted in 31.1% of the connections having 'poor', 57.4% 'fair' and 11.5% 'good' to 'excellent' reliability (median ICC=0.45, IQR=0.38-0.53). Leaving out the scrubbing preprocessing step (27.0% of the connections classified as having 'poor', 70.0% as having 'fair' and 3.0% as having 'good' to 'excellent' reliability, median ICC=0.44, IQR=0.40-0.49) or bandpass filter (23.1% of the connections classified as having 'poor', 73.5% 'fair' and 3.3% 'good' to 'excellent' reliability, median ICC=0.44, IQR=0.40-0.49) did not have a strong effect on the reconstruction reliability. Examining the reliability of all connections, also including connections below the proportional threshold in the group connectome, showed lower ICC scores indicating increased reliability after thresholding functional connectivity (Supplementary Figure 2).

Discussion

This software paper discusses CATO, a software toolbox for the reconstruction of structural and functional connectome maps. We describe the modular processing steps in the structural and functional pipelines and detail the implemented reconstruction methods, together with benchmark values to provide researchers insight in the sensitivity-specificity trade-off and reliability of connectome reconstructions.

Benchmarking structural reconstruction parameters and methods with respect to the ITC2015 challenge, shows that reconstruction methods in CATO have similar strengths and difficulties in reconstructing tracts as submissions to the ITC2015 challenge. ITC2015 datasets generally

show difficulty in correctly reconstructing cortico-spinal and commissural tracts, and so does CATO. The anterior commissure and posterior commissure, tracts that were not reconstructed were considered “very hard” in the ITC2015 challenge, likely due to their small cross-sectional diameter (< 2mm) and the left cortico-spinal tract was considered “hard” to reconstruct.

We suggest that the percentage of valid fibers in CATO reconstructions, which is now relatively low, could increase when reconstructed fiber clouds are further post-processed. The best performing DTI-based submissions of the ITC2015 challenge, for example, performed manual and semi-automatic editing to clean up the fiber cloud. Post-processing can be achieved manually, but also through automatic methods such as fiber clustering in which fiber bundles are detected and small fiber clusters excluded (Garyfallidis, Brett et al. 2012). Post-processing and post-selection on the resulting connectivity matrices may further improve connectome reconstruction, for example by applying group-mask matrix thresholding (van Wijk, Stam et al. 2010) (as we show here to improve ICC scores on the functional matrices) or for example by filtering out potential spurious connections based on anatomical priors (Schiavi, Ocampo-Pineda et al. 2020).

Benchmarking the sensitivity and specificity of the reconstructed structural connectomes showed that implemented reconstruction methods, at their default settings, have relatively high specificity compared to submissions of the ITC2015 challenge (Figure 3). Such strict connectome reconstructions are beneficial as false positive connections are estimated to have twice as much impact on network measures as false negatives (Zalesky, Fornito et al. 2016),

but have a negative effect on the level of sensitivity of finding more complex fiber bundles (i.e. at the price of having more false negatives). Exploring the effect of reconstruction parameters showed various ways in which researchers can tune the sensitivity and specificity profile of their reconstructions (Supplementary Figure 1). These benchmark findings suggest that the presented methods perform within the range of other powerful structural connectome reconstruction software tools (see Table 1) and provide a competitive connectome reconstruction toolset with the flexibility to tune the sensitivity and specificity trade-off of reconstructions.

The functional pipeline was benchmarked in a test-retest design based on repeated rs-fMRI data from the HCP dataset. Reconstructed functional connectivity showed reasonable reliability with around 70% of the connections in the group matrix having 'fair' or better reliability (i.e. ICC > 0.4). The median ICC scores showed around 0.40 which is on par with other functional methods that display an average ICC of individual connections of around 0.29 (Noble, Scheinost et al. 2019).

Methodological issues inherent to diffusion MRI and rs-fMRI have to be considered when using CATO. Diffusion MRI is limited by providing only an indirect measure of connectivity, with no information on the directionality of connections, having low spatial resolution and sensitivity to physiological motion (Mori and Zhang, 2006). Similarly, resting-state fMRI is limited by depending on correlations in the BOLD signal intensity, spatial and temporal resolution, and can be sensitive to physiological motion (Constable 2006). In our evaluation, we focused on benchmarking connectome reconstructions, rather than optimizing

performance, but a recent study suggests that careful selection of network reconstruction parameters and preprocessing steps can mitigate the impact of physical motion on structural connectivity (Oldham, Arnatkevičiūtė et al. 2020) and functional connectivity (Esteban, Markiewicz et al. 2019). Further, the calibration of reconstruction methods and parameters was based on the ITC2015 ground-truth white matter skeleton and the set of derived simulated diffusion images, but having only one data point for comparison does limit the generalizability of the results. The functional reconstruction pipeline was benchmarked in a test-retest design using the HCP dataset, but this design is likewise limited by the notion that high ICC coefficients can result from sensitivity to persistent covariates.

We hope that the modular organization, inclusion of both structural and functional connectivity reconstruction pipelines, and ease of use will be beneficial to researchers in the brain connectivity and neuroimaging field. Earlier versions of the toolbox have been used in previous studies examining connectome organization in the healthy brain (van den Heuvel, Scholtens et al. 2019), across various diseases (e.g. (Repple, Mauritz et al. 2020)), neonatal brain organization (Keunen, van der Burgh et al. 2018) and comparative connectomics (Ardesch, Scholtens et al. 2019). In addition, the availability of a diverse set of connectivity reconstruction packages enables the neuroscience community to replicate results using independent analysis pipelines. CATO is presented as an open-source connectivity reconstruction toolbox, shared under the MIT License, and will be further developed in dialog with users and other developers through our website and public GitHub repository.

Acknowledgements

Siemon C. de Lange was supported by the Amsterdam Neuroscience alliance grant. Martijn P. van den Heuvel was supported by an ALW open (ALWOP.179), a VIDI (452-16-015) grant from the Netherlands Organization for Scientific Research (NWO) and a Fellowship of MQ.

References

Alexander, A. L., J. E. Lee, M. Lazar and A. S. Field (2007). "Diffusion tensor imaging of the brain." Neurotherapeutics : the journal of the American Society for Experimental NeuroTherapeutics **4**(3): 316-329.

Andersson, J. L. R. and S. N. Sotiropoulos (2016). "An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging." NeuroImage **125**: 1063-1078.

Ardesch, D. J., L. H. Scholtens, L. Li, T. M. Preuss, J. K. Rilling and M. P. van den Heuvel (2019). "Evolutionary expansion of connectivity between multimodal association areas in the human brain compared with chimpanzees." Proc Natl Acad Sci U S A **116**(14): 7101-7106.

Bassett, D. S. and O. Sporns (2017). "Network neuroscience." Nat Neurosci **20**(3): 353-364.

Bullmore, E. and O. Sporns (2012). "The economy of brain network organization." Nature Reviews Neuroscience **13**(5): 336-336.

Cammoun, L., X. Gigandet, D. Meskaldji, J. P. Thiran, O. Sporns, K. Q. Do, P. Maeder, R. Meuli and P. Hagmann (2012). "Mapping the human connectome at multiple scales with diffusion spectrum MRI." Journal of Neuroscience Methods **203**(2): 386-397.

Chang, L.-C., D. K. Jones and C. Pierpaoli (2005). "RESTORE: Robust estimation of tensors by outlier rejection." Magnetic Resonance in Medicine **53**(5): 1088-1095.

Cicchetti, D. V. (1994). "Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology." Psychological assessment **6**(4): 284.

Constable, R. T. (2006). Challenges in fMRI and Its Limitations. Functional MRI: Basic Principles and Clinical Applications. S. H. Faro and F. B. Mohamed. New York, NY, Springer New York**:** 75-98.

de Reus, M. A. (2015). An eccentric perspective on brain networks, Uitgeverij BOXPress.

Desikan, R. S., F. Ségonne, B. Fischl, B. T. Quinn, B. C. Dickerson, D. Blacker, R. L. Buckner, A. M. Dale, R. P. Maguire, B. T. Hyman, M. S. Albert and R. J. Killiany (2006). "An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest." NeuroImage **31**(3): 968-980.

Esteban, O., C. J. Markiewicz, R. W. Blair, C. A. Moodie, A. I. Isik, A. Erramuzpe, J. D. Kent, M. Goncalves, E. DuPre, M. Snyder, H. Oya, S. S. Ghosh, J. Wright, J. Durnez, R. A. Poldrack and K. J. Gorgolewski (2019). "fMRIPrep: a robust preprocessing pipeline for functional MRI." Nature Methods **16**(1): 111-116.

Fischl, B. (2012). "Freesurfer." Neuroimage **62**(2): 774-781.

Fischl, B., A. Van Der Kouwe, C. Destrieux, E. Halgren, F. Ségonne, D. H. Salat, E. Busa, L. J. Seidman, J. Goldstein, D. Kennedy, V. Caviness, N. Makris, B. Rosen and A. M. Dale

(2004). "Automatically Parcellating the Human Cerebral Cortex." Cerebral Cortex **14**(1): 11-22.
Fornito, A., A. Zalesky and M. Breakspear (2015). "The connectomics of brain disorders." Nature Reviews Neuroscience **16**(3): 159-172.
Fox, M. D., D. Zhang, A. Z. Snyder and M. E. Raichle (2009). "The global signal and observed anticorrelated resting state brain networks." J Neurophysiol **101**(6): 3270-3283.
Garyfallidis, E., M. Brett, B. Amirbekian, A. Rokem, S. Van Der Walt, M. Descoteaux and I. Nimmo-Smith (2014). "Dipy, a library for the analysis of diffusion MRI data." Frontiers in neuroinformatics **8**: 8.
Garyfallidis, E., M. Brett, M. M. Correia, G. B. Williams and I. Nimmo-Smith (2012). "Quickbundles, a method for tractography simplification." Frontiers in neuroscience **6**: 175.
Gavin, H. P. (2019). "The Levenberg-Marquardt algorithm for nonlinear least squares curve-fitting problems." http://people.duke.edu/~hpgavin/ce281/lm.pdf Accessed 26 November 2019 2019.
Glasser, M. F., S. N. Sotiropoulos, J. A. Wilson, T. S. Coalson, B. Fischl, J. L. Andersson, J. Xu, S. Jbabdi, M. Webster, J. R. Polimeni, D. C. Van Essen and M. Jenkinson (2013). "The minimal preprocessing pipelines for the Human Connectome Project." NeuroImage **80**: 105-124.
Greve, D. N. and B. Fischl (2009). "Accurate and robust brain image alignment using boundary-based registration." Neuroimage **48**(1): 63-72.
Hagmann, P., L. Cammoun, X. Gigandet, R. Meuli, C. J. Honey, J. Van Wedeen, O. Sporns and V. J. Wedeen (2008). "Mapping the Structural Core of Human Cerebral Cortex." PLoS Biol **6**(7): e159-e159.
Jenkinson, M., P. Bannister, M. Brady and S. Smith (2002). "Improved optimization for the robust and accurate linear registration and motion correction of brain images." Neuroimage **17**(2): 825-841.
Jenkinson, M., C. F. Beckmann, T. E. Behrens, M. W. Woolrich and S. M. Smith (2012). "Fsl." Neuroimage **62**(2): 782-790.
Jones, D. K. (2004). "The Effect of Gradient Sampling Schemes on Measures Derived From Diffusion Tensor MRI: A Monte Carlo Study." Magnetic Resonance in Medicine **51**: 807-815.
Keunen, K., H. K. van der Burgh, M. A. de Reus, P. Moeskops, R. Schmidt, L. J. Stolwijk, S. C. de Lange, I. Isgum, L. S. de Vries, M. J. Benders and M. P. van den Heuvel (2018). "Early human brain development: insights into macroscale connectome wiring." Pediatr Res **84**(6): 829-836.
Leemans, A. and D. K. Jones (2009). "The B-matrix must be rotated when correcting for subject motion in DTI data." Magn Reson Med **61**(6): 1336-1349.
Maier-Hein, K. H., P. F. Neher, J.-C. Houde, M.-A. Côté, E. Garyfallidis, J. Zhong, M. Chamberland, F.-C. Yeh, Y.-C. Lin, Q. Ji, W. E. Reddick, J. O. Glass, D. Q. Chen, Y. Feng, C. Gao, Y. Wu, J. Ma, H. Renjie, Q. Li, C.-F. Westin, S. Deslauriers-Gauthier, J. O. O. González, M. Paquette, S. St-Jean, G. Girard, F. Rheault, J. Sidhu, C. M. W. Tax, F. Guo, H. Y. Mesri, S. Dávid, M. Froeling, A. M. Heemskerk, A. Leemans, A. Boré, B. Pinsard, C. Bedetti, M. Desrosiers, S. Brambati, J. Doyon, A. Sarica, R. Vasta, A. Cerasa, A. Quattrone, J. Yeatman, A. R. Khan, W. Hodges, S. Alexander, D. Romascano, M. Barakovic, A. Auría, O. Esteban, A. Lemkaddem, J.-P. Thiran, H. E. Cetingul, B. L. Odry, B. Mailhe, M. S. Nadar, F. Pizzagalli, G. Prasad, J. E. Villalon-Reina, J. Galvis, P. M. Thompson, F. D. S. Requejo, P. L. Laguna, L. M. Lacerda, R. Barrett, F. Dell'Acqua, M. Catani, L. Petit, E. Caruyer, A. Daducci, T. B. Dyrby, T. Holland-Letz, C. C. Hilgetag, B. Stieltjes and M. Descoteaux (2017). "The challenge of

mapping the human connectome based on diffusion tractography." Nature Communications **8**(1): 1349-1349.
McGraw, K. O. and S. P. Wong (1996). "Forming inferences about some intraclass correlation coefficients." Psychological methods **1**(1): 30.
Mennes, M., B. B. Biswal, F. X. Castellanos and M. P. Milham (2013). "Making data sharing work: the FCP/INDI experience." Neuroimage **82**: 683-691.
Mori, S., B. J. Crain, V. P. Chacko and P. C. M. van Zijl (1999). "Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging." Annals of Neurology **45**(2): 265-269.
National Academies of Sciences, E. and Medicine (2019). Reproducibility and Replicability in Science. Washington, DC, The National Academies Press.
Noble, S., D. Scheinost and R. T. Constable (2019). "A decade of test-retest reliability of functional connectivity: A systematic review and meta-analysis." NeuroImage **203**: 116157.
Oldham, S., A. Arnatkevičiūtė, R. E. Smith, J. Tiego, M. A. Bellgrove and A. Fornito (2020). "The efficacy of different preprocessing steps in reducing motion-related confounds in diffusion MRI connectomics." NeuroImage **222**: 117252.
Power, J. D., K. a. Barnes, A. Z. Snyder, B. L. Schlaggar and S. E. Petersen (2012). "Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion." NeuroImage **59**(3): 2142-2154.
Repple, J., M. Mauritz, S. Meinert, S. C. de Lange, D. Grotegerd, N. Opel, R. Redlich, T. Hahn, K. Förster and E. J. Leehr (2020). "Severity of current depression and remission status are associated with structural connectome alterations in major depressive disorder." Molecular psychiatry **25**(7): 1550-1558.
Rubinov, M. and O. Sporns (2010). "Complex network measures of brain connectivity: uses and interpretations." NeuroImage **52**(3): 1059-1069.
Schiavi, S., M. Ocampo-Pineda, M. Barakovic, L. Petit, M. Descoteaux, J.-P. Thiran and A. Daducci (2020). "A new method for accurate in vivo mapping of human brain connections using microstructural and anatomical information." Science Advances **6**(31): eaba8245.
Scholtens, L. H., M. A. de Reus, S. C. de Lange, R. Schmidt and M. P. van den Heuvel (2016). "An MRI Von Economo – Koskinas atlas." NeuroImage.
Taylor, P. A. and Z. S. Saad (2013). "FATCAT: (an efficient) Functional and Tractographic Connectivity Analysis Toolbox." Brain connectivity **3**(5): 523-535.
Tournier, J. D., F. Calamante and A. Connelly (2007). "Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution." NeuroImage **35**(4): 1459-1472.
Tuch, D. S. (2004). "Q-ball imaging." Magn Reson Med **52**(6): 1358-1372.
van den Heuvel, M. P., L. H. Scholtens, H. K. van der Burgh, F. Agosta, C. Alloza, C. Arango, B. Auyeung, S. Baron-Cohen, S. Basaia, M. J. N. L. Benders, F. Beyer, L. Booij, K. P. J. Braun, G. B. Filho, W. Cahn, D. M. Cannon, T. M. Chaim-Avancini, S. S. M. Chan, E. Y. H. Chen, B. Crespo-Facorro, E. A. Crone, U. Dannlowski, S. M. C. de Zwarte, B. Dietsche, G. Donohoe, S. D. Plessis, S. Durston, C. M. Díaz-Caneja, A. M. Díaz-Zuluaga, R. Emsley, M. Filippi, T. Frodl, M. Gorges, B. Graff, D. Grotegerd, D. Gąsecki, J. M. Hall, L. Holleran, R. Holt, H. J. Hopman, A. Jansen, J. Janssen, K. Jodzio, L. Jäncke, V. G. Kaleda, J. Kassubek, S. K. Masouleh, T. Kircher, M. G. J. C. Koevoets, V. S. Kostic, A. Krug, S. M. Lawrie, I. S. Lebedeva, E. H. M. Lee, T. A. Lett, S. J. G. Lewis, F. Liem, M. V. Lombardo, C. Lopez-Jaramillo, D. S. Margulies, S. Markett, P. Marques, I. Martínez-Zalacaín, C. McDonald, A. M. McIntosh, G. McPhilemy, S. L. Meinert, J. M. Menchón, C. Montag, P. S. Moreira, P. Morgado, D. O. Mothersill, S.

Mérillat, H.-P. Müller, L. Nabulsi, P. Najt, K. Narkiewicz, P. Naumczyk, B. Oranje, V. Ortiz-Garcia de la Foz, J. S. Peper, J. A. Pineda, P. E. Rasser, R. Redlich, J. Repple, M. Reuter, P. G. P. Rosa, A. N. V. Ruigrok, A. Sabisz, U. Schall, S. Seedat, M. H. Serpa, S. Skouras, C. Soriano-Mas, N. Sousa, E. Szurowska, A. S. Tomyshev, D. Tordesillas-Gutierrez, S. L. Valk, L. H. van den Berg, T. G. M. van Erp, N. E. M. van Haren, J. M. C. van Leeuwen, A. Villringer, C. H. Vinkers, C. Vollmar, L. Waller, H. Walter, H. C. Whalley, M. Witkowska, A. V. Witte, M. V. Zanetti, R. Zhang and S. C. d. Lange (2019). "10Kin1day: A Bottom-Up Neuroimaging Initiative." Frontiers in Neurology **10**(425).

van Wijk, B. C. M., C. J. Stam and A. Daffertshofer (2010). "Comparing brain networks of different size and connectivity density using graph theory." PloS one **5**(10): e13701-e13701.

Yeh, F.-C., V. J. Wedeen and W.-Y. I. Tseng (2010). "Generalized q-sampling imaging." IEEE transactions on medical imaging **29**(9): 1626-1635.

Zalesky, A., A. Fornito, L. Cocchi, L. L. Gollo, M. P. van den Heuvel and M. Breakspear (2016). "Connectome sensitivity or specificity: which is more important?" NeuroImage **142**: 407-420.

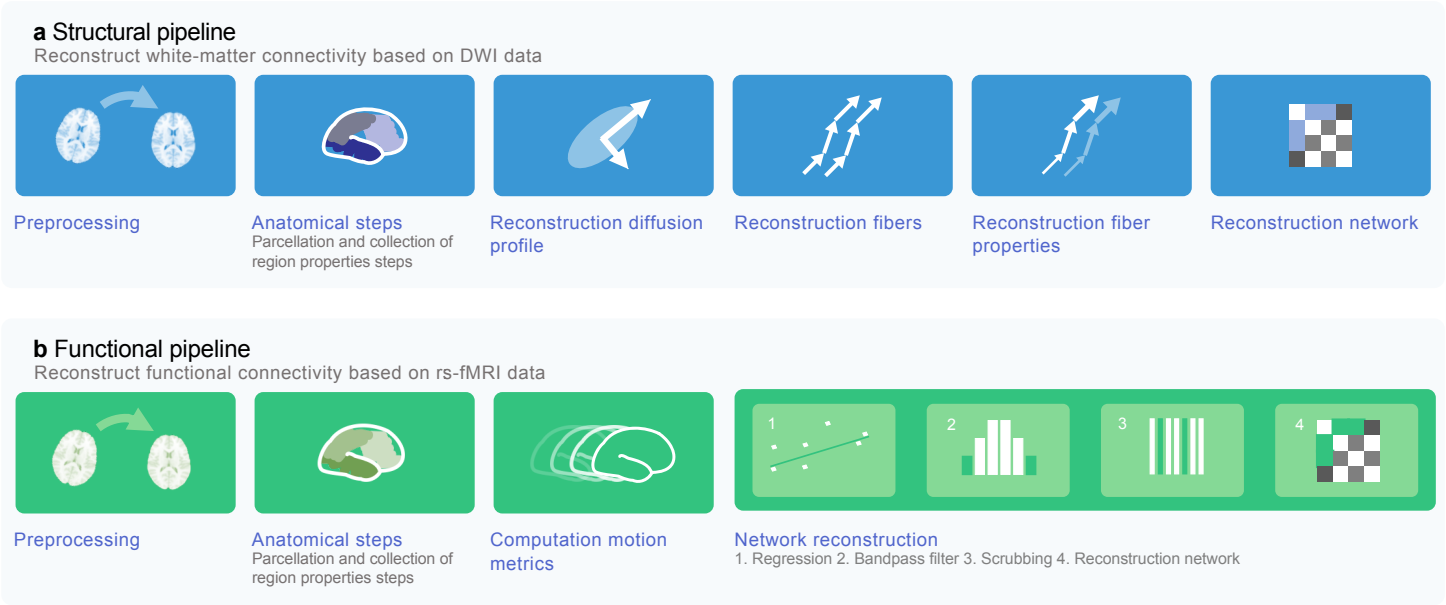

Figure 1. Overview CATO toolbox and pipelines. CATO is a modular toolbox consisting of two pipelines for reconstruction of respectively structural (blue) and functional (green) connectivity. Each pipeline consists of multiple separate processing steps, with the anatomical parcellation steps being shared between pipelines.

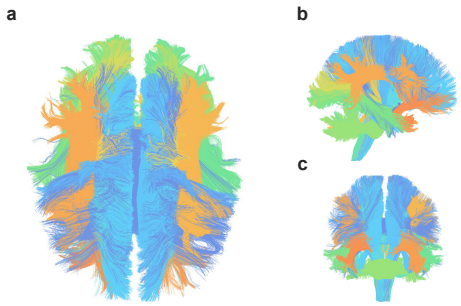

Figure 2. Ground truth white matter skeleton of the ISMRM 2015 Tractography challenge. Fiber cloud is shown from top (a), left-side (b) and front (c).

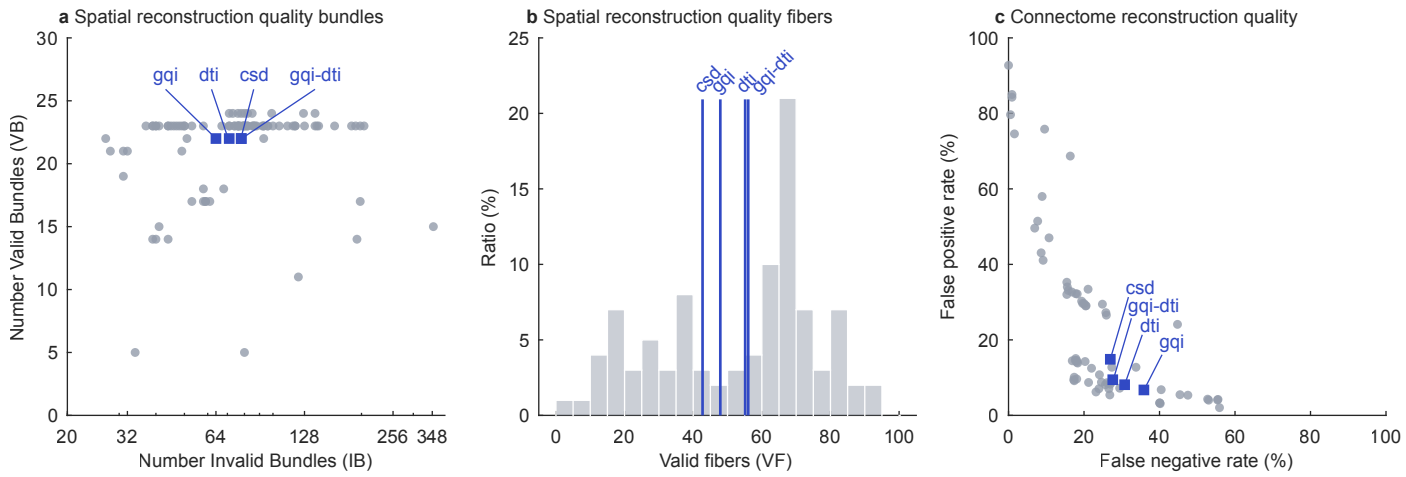

Figure 3. Benchmarking reconstructions. The spatial reconstruction quality of fiber cloud reconstructions obtained using CATO were benchmarked with respect to different reconstruction settings, and submissions in the ITC2015 challenge. (a) The spatial reconstruction quality of fibers was assessed by the number of valid and invalid reconstructed bundles. The results of CATO are indicated by purple squares and results from submissions to the ITC2015 are indicated by grey circles. Reconstructions showed a similar number of valid bundles (22) and invalid bundles (64 - 78) as ITC2015 submissions (median VB = 23, median IB = 77). (b) The spatial reconstruction quality was also assessed by the percentage of valid reconstructed fibers. The results of CATO are presented by purple lines and the grey histogram shows the valid fiber distribution across ITC2015 submissions. The percentage of valid fibers in the reconstructions ranged between 43% and 55% and ITC2015 submissions showing percentages between 3.75% and 92.5% (medianVF = 62.6%). (c) Connectome reconstruction quality (aim of CATO) was assessed by the level of false negative and false positive connection reconstructions. CATO showed relatively high specificity (false positive rates: 6.7% - 14.9%; and false negative rates: 27.0% - 35.9%) compared to the ITC2015 submissions (median false positive rates: 14.5%; and median false negative rates: 20.3%).

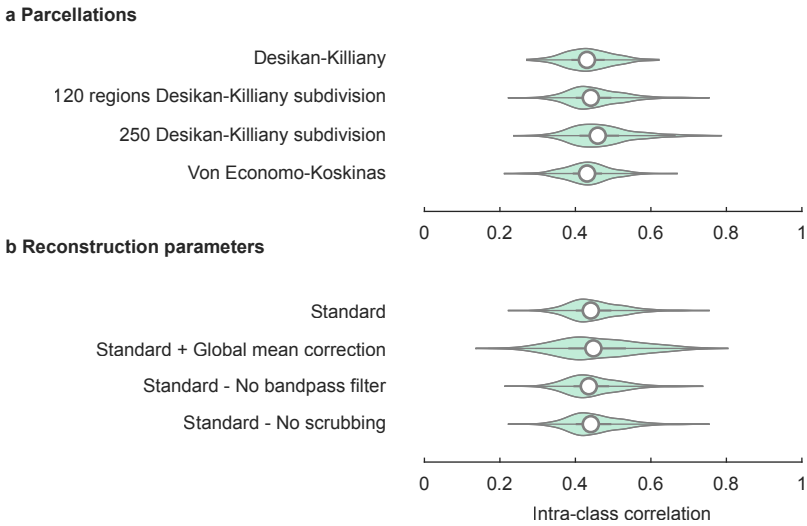

Figure 4. Functional pipeline test-retest reliability. (a) Test-retest reliability of functional connections for the Desikan-Killiany atlas (median ICC=0.43, IQR=0.39-0.48), the 120 regions sub-parcellation of the Desikan-Killiany atlas (median ICC=0.44, IQR=0.40-0.49), 250 regions sub-parcellation of the Desikan Killiany atlas (median ICC=0.46, IQR=0.41-0.51) and the Von Economo-Koskinas atlas (median ICC=0.43, IQR=0.39-0.47). (b) Test-retest reliability of functional connections of the Desikan-Killiany atlas, using standard HCP-default parameters, standard processing with additional global mean regression (median ICC=0.45, IQR=0.38-0.53), standard processing excluding bandpass filtering (median ICC=0.44, IQR=0.40-0.49) and standard processing excluding scrubbing (median ICC=0.44, IQR=0.40-0.49). Across all parcellations and reconstruction methods, around half of the connections had ICC > 0.4, indicating 'fair' reliability. Boxes indicate the interval between the 25th and 75th percentiles (quartile q_1 and q_3), whiskers indicate the interval between q_1 − 1.5 × (q_3-q_1) and q_3 + 1.5 × (q_3-q_1), the white circles indicate median values.

Tables

Name	Reference	Modality	Programming language
ANFI (FATCAT)	(Taylor and Saad 2013)	Structural and Functional	C and R
BrainCAT	(Marques, Soares et al. 2013)	Structural and Functional	Objective-C
Dipy	(Garyfallidis, Brett et al. 2014)	Structural	Python
DSI studio	http://dsi-studio.labsolver.org	Structural	C++
ExploreDTI	(Leemans, Jeurissen et al. 2009)	Structural	MATLAB
FSL	(Jenkinson, Beckmann et al. 2012)	Structural and Functional	MATLAB and C
MRtrix	(Tournier, Smith et al. 2019)	Structural	Python and C++
SPM	https://www.fil.ion.ucl.ac.uk/spm	Functional	MATLAB and C
TrackVis and Diffusion Toolkit	(Wang, Benner et al. 2007)	Structural	C++
TractoFlow	(Theaud, Houde et al. 2019)	Structural	Nextflow and Singularity

Table 1. Overview of connectome reconstruction software packages.

Abbreviation	Full name
CA	Anterior commissure
CC	Corpus callosum
CG - left	Cingulum - left
CG - right	Cingulum - right
CP	Posterior commissure
CST - left	Cortico-spinal tract - left
CST - right	Cortico-spinal tract - right
Fornix	Fornix
FPT - left	Frontopontine tracts - left
FPT - right	Frontopontine tracts - right
ICP - left	Inferior cerebellar peduncle - left

ICP - right	Inferior cerebellar peduncle - right
ILF - left	Inferior longitudinal fasciculus - left
ILF - right	Inferior longitudinal fasciculus - right
MCP	Middle cerebellar peduncle
OR - left	Optic radiation - left
OR - right	Optic radiation - right
POPT - left	parieto-occipital pontine tract - left
POPT - right	parieto-occipital pontine tract - right
SCP - left	Superior cerebellar peduncle - left
SCP - right	Superior cerebellar peduncle - right
SLF - left	Superior longitudinal fasciculus - left
SLF - right	Superior longitudinal fasciculus - right
UF - left	Uncinate fasciculus - left
UF - right	Uncinate fasciculus - right

Table 2. Overview of white matter bundles in ground truth white matter skeleton in the ISMRM 2015 Tractography challenge.