Published in final edited form as:
Med Image Comput Comput Assist Interv. 2013 ; 16(0 1): 647–654.

On Describing Human White Matter Anatomy: The White Matter Query Language

Demian Wassermann[1,*], **Nikos Makris**[2], **Yogesh Rathi**[1], **Martha Shenton**[1], **Ron Kikinis**[1], **Marek Kubicki**[1], and **Carl-Fredrik Westin**[1]

Demian Wassermann: demian@bwh.harvard.edu

[1]Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA

[2]Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

Abstract

The main contribution of this work is the careful syntactical definition of major white matter tracts in the human brain based on a neuroanatomist's expert knowledge. We present a technique to formally describe white matter tracts and to automatically extract them from diffusion MRI data. The framework is based on a novel query language with a near-to-English textual syntax. This query language allows us to construct a dictionary of anatomical definitions describing white matter tracts. The definitions include adjacent gray and white matter regions, and rules for spatial relations. This enables automated coherent labeling of white matter anatomy across subjects. We use our method to encode anatomical knowledge in human white matter describing 10 association and 8 projection tracts per hemisphere and 7 commissural tracts. The technique is shown to be comparable in accuracy to manual labeling. We present results applying this framework to create a white matter atlas from 77 healthy subjects, and we use this atlas in a proof-of-concept study to detect tract changes specific to schizophrenia.

1 Introduction

Diffusion magnetic resonance imaging (dMRI) is a technique that allows to probe the structure of white matter the human brain *in vivo*. dMRI streamline tractography has provided the opportunity for non-invasive investigation of white matter anatomy. Most common methods for isolating fiber bundles based on streamline tractography require the manual placement of multiple regions of interest (ROIs). These methods include an approach that starts from seed points within a predefined region of interest, and then calculates and preserves only tracts that touch other predefined ROIs [1]. A different approach creates seed points throughout the entire brain (whole brain tractography) keeping tracts that pass through conjunctions, disjunctions or exclusions of ROIs either off-line [2] or interactively [3]. An alternative method to manual placement of ROIs is to use a clustering approach [4–6]. Clustering methods are in general fully automatic, unguided, and take advantage of the similarity of fiber paths. However, incorporating precise information about human anatomy into a clustering method is difficult.

[*]Camera ready version for the Medical Image Computing and Computed Assisted Intervention 2013 conference

The ability to target specific tracts for analyses compared to whole brain studies increases the statistical power and sensitivity of the study, and simplifies the interpretation of results, but requires precise consistent delineation of the tracts across subjects. Although several fascicules, such as the cingulum bundle, are widely recognized and well defined in field of neuroanatomy, there are others which existence or subdivisions is still a matter of discussion. Examples are the different systems to define and subdivide the superior longitudinal fasciculus (SLF) proposed by Catani [7] and Makris [8], and the discussion regarding existence of the inferior occipito-frontal fasciculus (IOFF) in humans [9]. Three major challenges make it difficult to extend and reproduce tractography-based dissections and tract-specific analyses: 1) the anatomical dissection of the white matter into specific fascicles is currently in discussion; 2) comparing fascicle definitions across atlases is a difficult task due to the lack of a system to unify the definitions; 3) semi-automated approaches are usually based on a fixed set of fascicles difficult to extend due to amount of technical knowledge needed for the task.

The main contribution of this paper is a system to express anatomical descriptions of white matter tracts in a near-to-English textual language, and we believe this work can help address the above challenges. We designed this textual language to be human-readable to make the tract descriptions easy to change and extend without the need of an engineering background and to be used for automated white matter virtual dissections. The white matter query language (WMQL) proposed in this paper has several applications. For example Wakana's and Catani's definition of tracts using regions of interest (ROIs) [10, 2], can readily be represented using WMQL definitions. This will yield an human readable definition that also can be extended by an anatomist for finer division. Another interesting application is to post-process clustering results and automatically label clusters as anatomically known tracts [4–6]. In this paper we present definitions of different tracts from current literature and formulated WMQL descriptions. Our anatomy dictionary currently contains descriptions of 10 association and 8 projection tracts per hemisphere, and 7 commissural tracts. An implementation of WMQL as well as the definitions specified in this publication can be downloaded at http://demianw.github.com/tract_querier

2 Methods

We designed the queries in white matter query language (WMQL) to formalize current descriptions in anatomy literature. Such descriptions are constructed in terms of different relationships between gyri, sulci, subcortical structures or white matter areas.The operations of WMQL, can be divided into 3 groups: 1) **anatomical terms** stating if a tract traverses or ends in a certain brain structure, 2) **relative position terms** indicating whether the tracts are, for instance, medial or frontal to a structure like the amygdala, and 3) **logical operations** like conjunction, disjunction or exclusion of the previous two types of clauses. We illustrate these three types of operations in fig. 1.

To apply WMQL queries to dissect dMRI-based tractographies, like in fig. 1, we need to situate gyri, sulci and other structures used for the queries relative to the tractography. For this, we overlay an atlas of the brain structures on top the tractography, in this work we used the cortical and white matter parcellation based on the Desikan atlas as described by [11]

and the neuroanatomic structure segmentation by [12] which are readily available in FreeSurfer (http://www.freesurfer.org). We provide the details of this process in the section MRI Data Acquisition and Processing. It is worth noting that WMQL does not depend on a particular atlas.

To implement the white matter query language, we group the tracts in sets representing whether they have an endpoint in each anatomical label (fig. 1b); traverse anatomical label (fig. 1c); or their position relative to each anatomical label (fig. 1d–f). Then, for each anatomical label, like the amygdala, we have six sets:

1. endpoints_in(amygdala): all tracts with at least an endpoint in the amygdala.
2. amygdala: all the tracts traversing the amygdala.
3. anterior_of(amygdala), posterior_of(amygdala), medial_of(amygdala), lateral_of(amygdala), superior_of(amygdala), inferior_of(amygdala) : containing the tracts traversing brain areas delimited by their relative position to the amygdala

We calculate the endpoint sets (1) by computing which label is the tract touching at each endpoint and adding the tract to the corresponding set. The sets representing label traversals (2) are computed by following each tract and adding it to the set corresponding to a particular label, like the amygdala, if tract traverses it. Finally, we compute the 6 relative positioning sets (3) by checking the points the tract traverses with respect to a label. For instance, every tract that traverses the area anterior to the amygdala, shown in orange in fig. 1e, will be added to the set anterior_of(amygdala). We implement this efficiently using a tree of axis-aligned bounding boxes of the structures as a spatial index of the labels and then computing the relative positions of the tracts with respect to the bounding boxes composing each label [13]. For two sets of tracts a and b and the set of all tracts L, we formalize the WQML logical operations as follows:

$$\text{a } \mathbf{or} \text{ b} := \{tract : tract \in a \cup b\} \qquad \mathbf{only}(\text{a}) := \{tract : tract \in a \wedge tract \notin (L \backslash a)\}$$
$$\text{a } \mathbf{and} \text{ b} := \{tract : tract \in a \cap b\} \qquad \text{a } \mathbf{not\ in} \text{ b} := \{tract : tract \in a \wedge tract \notin b\}$$

We have illustrated these operations in fig. 1 a), c); and g)-i). Tracts in WMQL are defined by using an assignment operation, for instance, we defined the left UF in fig. 1i as assigning a meaning to UF.left:

```
UF.left
= insula.left and
(lateral_frontal.left or
medial_frontal.left or
orbito_frontal.left)
and endpoints_in(temporal.left and anterior_of(amygdala.left))
not in
hemisphere.right
```

To simplify the definitions of tracts in both hemispheres with a single assignment, we included the suffixes.side and.opposite to WMQL. Creating a bihemispheric definition becomes:

```
UF.side
= insula.side and
(lateral_frontal.side or
medial_frontal.side or
orbito_frontal.side)
and
endpoints_in(temporal.side and
anterior_of(amygdala.side))
not in
hemisphere.opposite
```

In this way WMQL allows a single definition for tracts found in both hemispheres simultaneously. The formalization of WMQL as the basic set operations allows us to define white matter tracts using all the flexibility and expressiveness power of set theory and propositional logic. Finally, to implement the WMQL language, we defined the WMQL in Backus normal form; we used an LL(1) parser for this grammar which transforms WMQL expressions into an abstract syntax tree; and we implemented an algorithm to traverse the tree evaluating the WMQL operations.

3 Results

3.1 Development of WMQL

We are using WMQL to formalize tract descriptions from classic neuroanatomy textbooks and current literature on the anatomy of the white matter [14, 10, 8, 15]. In table 1, we show the queries for 10 association tracts. In the following, for each tract we describe in WMQL, we provide a description derived from anatomy literature and the derived WMQL query. The population results of these queries are shown in fig. 2 along with the commisural and projection tracts whose definitions we did not include for space reasons. Using the WMQL, we have constructed a comprehensive atlas based on high angular MRI (HARDI) tractography [16]. The combination of WMQL with current tractography technologies has enabled us to generalize previous atlases [10, 2] and extend them with white matter tracts not included in these like the middle longitudinal fasciculus (MdLF) or the three different superior longitudinal fasciculi. In total, our atlas includes 10 long association tracts; 8 projection tracts per hemisphere (7 cortico-striatal and the Cortico-spinal tract) and the 7 sections of the corpus callosum according to [15].

3.2 Application of WMQL to Tract Extraction and Statistical Analyses

Diffusion-weighted images (DWI) from 77 healthy subjects (HS) (32.9 ± 12.4 years old of age; 64 males; right handed) and 20 male schizophrenic subjects (SZ) paired with the HS were acquired. DWI data were acquired on a GE Signa HDxt 3.0T (51 directions with b=900

s/mm2, 8 b=0 s/mm2 images, 1.7 mm^3 isotropic voxels). A T1 MRI acquisition was also performed (25.6cm2 field of view, $1mm^3$ isotropic voxels). For each DWI image, we obtained a 2-tensor full-brain tractography placing ten seeds per voxel and obtaining an average of one million tracts per subject. We overlaid a parcellation of the cortical and sub-cortical structures and the white matter on DWI images by processing the T1 images of each subject using FreeSurfer and registering the results to the DWI images using ANTS [17]. For each subject, this resulted in the subcortical structures labeled and the cortex the white matter parcellated [11].

Validation—To validate our extraction protocol, 2 different experts segmented 5 tracts (IOFF, UF, ILF, CST and AF) in each hemisphere of 10 healthy subjects following the protocol in [10]. Overlap between those and the ones extracted was calculated with the kappa measure [2]. For all tracts $k > .7$ which is considered good agreement, the worse being the left AF ($k = .71$) and the best the left CST ($k = .89$).

White Matter Atlas Generation with WMQL—For each subject we extracted 37 white matter tracts using WMQL queries derived from classical and current literature of the human brain white matter. Some of these are detailed in table 1. Processing each full brain tractography took 5 ± .25 min. to initialize and 7 ± .02 sec. per query.

To generate a tract atlas, we normalized all the FA maps to MNI space, created a population FA template using ANTS and then generated group effect maps for each tract [7]. To obtain the group effect maps, we started by calculating a binary visitation map for each tract of each subject. This map is a mask in MNI space where a voxel has a value of one if the tract traverses that voxel and 0 if it doesn't. We create the group effect map for each tract through voxel-wise statistics. The group effect map assesses the probability that a tract traverses each voxel. For this, we rejected the null hypothesis that that voxel is not traversed by such tract. We first smoothed the visitation maps with a 2mm (FWHM) isotropic Gaussian to approximate the distribution of the data to a Gaussian one [18]. Then, we rejected the hypothesis that the voxel does not belong to the tract, i.e. the mean traversal value over all subjects on that voxel is different to 0, by using a voxel wise t-test for a one-sample mean. We calculated the corrected significance using permutation testing (10,000 iterations) to avoid a high dependence on Gaussian assumptions [19]. We set the significance threshold for considering that a voxel belongs to the tract to p-value < 0.0001 corrected for multiple comparisons.

Group Differences in Schizophrenia—As a proof of concept, we analyzed the tracts that are not included in other atlases, but found in our WM atlas only: MdLF, SLF I, II and III in 20 SZ subjects and controls extracted from the HS paired by age and gender. We used the group maps to obtain a weighted average of FA resulting in one value per subject and performed a t-test for each tract correcting for age. We show these results in fig. 3. Results show agreement with recent discoveries in SZ in the MdLF tract [20]

4 Conclusion

In this work we have introduced the White Matter Query Language: a tool to represent anatomical knowledge of white matter tracts and extract them from full-brain tractographies. This tool is a complement to current semi-automated approaches in tract extraction based in clustering and can be used to implement ROI-based approaches. The utility of the tool was illustrated by constructing an atlas of white matter tracts from 77 healthy subjects and then performing a statistical analysis in schizophrenia. The textual descriptions in WMQL add transparency between the conceptualization of white matter fiber tracts and their extraction from a dMRI.

Acknowledgments

This work has been supported by NIH grants: R01MH074794, R01MH092862, P41RR013218, R01MH097979, P41EB015902 and Swedish Research Council (VR) grant 2012-3682.

References

1. Mori S, van Zijl PCM. Fiber tracking: principles and strategies-a technical review. NMR in Biomedicine. 2002
2. Wakana S, Caprihan A, Panzenboeck MM, Fallon JH, Perry M, Gollub RL, Hua K, Zhang J, Jiang H, Dubey P, Blitz A, van Zijl PC, Mori S. Reproducibility of quantitative tractography methods applied to cerebral white matter. NImg. 2007
3. Akers D, Sherbondy A, Mackenzie R, Dougherty R, Wandell B. Exploration of the brain's white matter pathways with dynamic queries. IEEE Viz. 2004
4. Wassermann D, Bloy L, Kanterakis E, Verma R, Deriche R. Unsupervised white matter fiber clustering and tract probability map generation: Applications of a Gaussian process framework for white matter fibers. NImg. 2010
5. O'Donnell LJ, Westin CF. Automatic Tractography Segmentation Using a High-Dimensional White Matter Atlas. TMI. 2007
6. Wang X, Grimson WEL, Westin CF. Tractography segmentation using a hierarchical Dirichlet processes mixture model. NImg. 2011
7. Thiebaut de Schotten M, Ffytche DH, Bizzi A, Dell'acqua F, Allin M, Walshe M, Murray R, Williams SC, Murphy DG, Catani M. Atlasing location, asymmetry and inter-subject variability of white matter tracts in the human brain with MR diffusion tractography. NImg. 2010
8. Makris N, Kennedy DN, McInerney S, Sorensen AG, Wang R, Caviness VS, Pandya D. Segmentation of Subcomponents within the Superior Longitudinal Fascicle in Humans: A Quantitative, In Vivo, DT-MRI Study. Cerebral Cortex. 2005
9. Schmahmann JD, Pandya D. The Complex History of the Fronto-Occipital Fasciculus. Journal of the History of the Neurosciences. 2007
10. Catani M, Thiebaut de Schotten M. A diffusion tensor imaging tractography atlas for virtual in vivo dissections. Cortex. 2008
11. Salat DH, Greve DN, Pacheco JL, Quinn BT, Helmer KG, Buckner RL, Fischl B. Regional white matter volume differences in nondemented aging and Alzheimer's disease. NImg. 2009
12. Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, van der Kouwe A, Killiany R, Kennedy D, Klaveness S, Montillo A, Makris N, Rosen B, Dale AM. Whole Brain Segmentation: Automated Labeling of Neuroanatomical Structures in the Human Brain. Neuron. 2002
13. Bergen, Gvd. Efficient Collision Detection of Complex Deformable Models using AABB Trees. Journal of Graphic Tools. 1997
14. Parent, A. Carpenter's Human Neuroanatomy. Williams & Wilkins; 1996.

15. Witelson SF, Kigar DL. Anatomical development of the corpus callosum in humans: A review with reference to sex and cognition. Brain Lateralization in children: Developmental implications. 1988
16. Malcolm JG, Michailovich O, Bouix S, Westin CF, Shenton ME, Rathi Y. A filtered approach to neural tractography using the Watson directional function. TMI. 2010
17. Avants B, Epstein C, Grossman M, Gee JC. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. MIA. 2008
18. Ashburner J, Friston KJ. Voxel-Based Morphometry—The Methods. NImg. 2000
19. Nichols TE, Holmes AP. Nonparametric permutation tests for functional neuroimaging: a primer with examples. HBM. 2002
20. Asami T, Saito Y, Whitford TJ, Makris N, Niznikiewicz M, McCarley RW, Shenton ME, Kubicki M. Abnormalities of middle longitudinal fascicle and disorganization in patients with schizophrenia. Schizophrenia Research. 2013

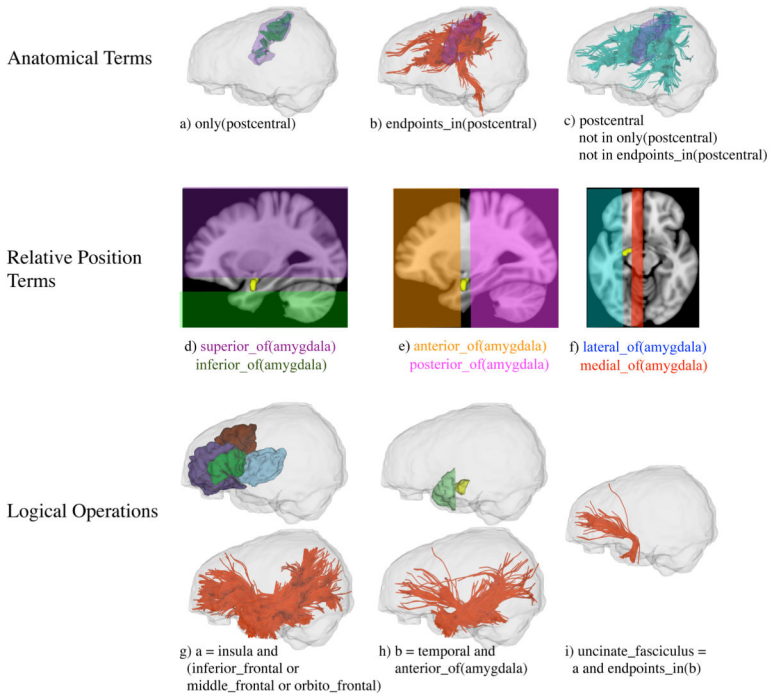

Fig. 1.

WMQL Terms (a–f) along with an example construction of a WMQL query (g–i). Regions in (g–i): insula (cyan); the orbito-frontal (purple), middle-frontal (brown) and inferior-frontal (dark green) gyri. h) shows the anterior temporal lobe (light green) defined as the section of the temporal lobe anterior to the amygdala (yellow).

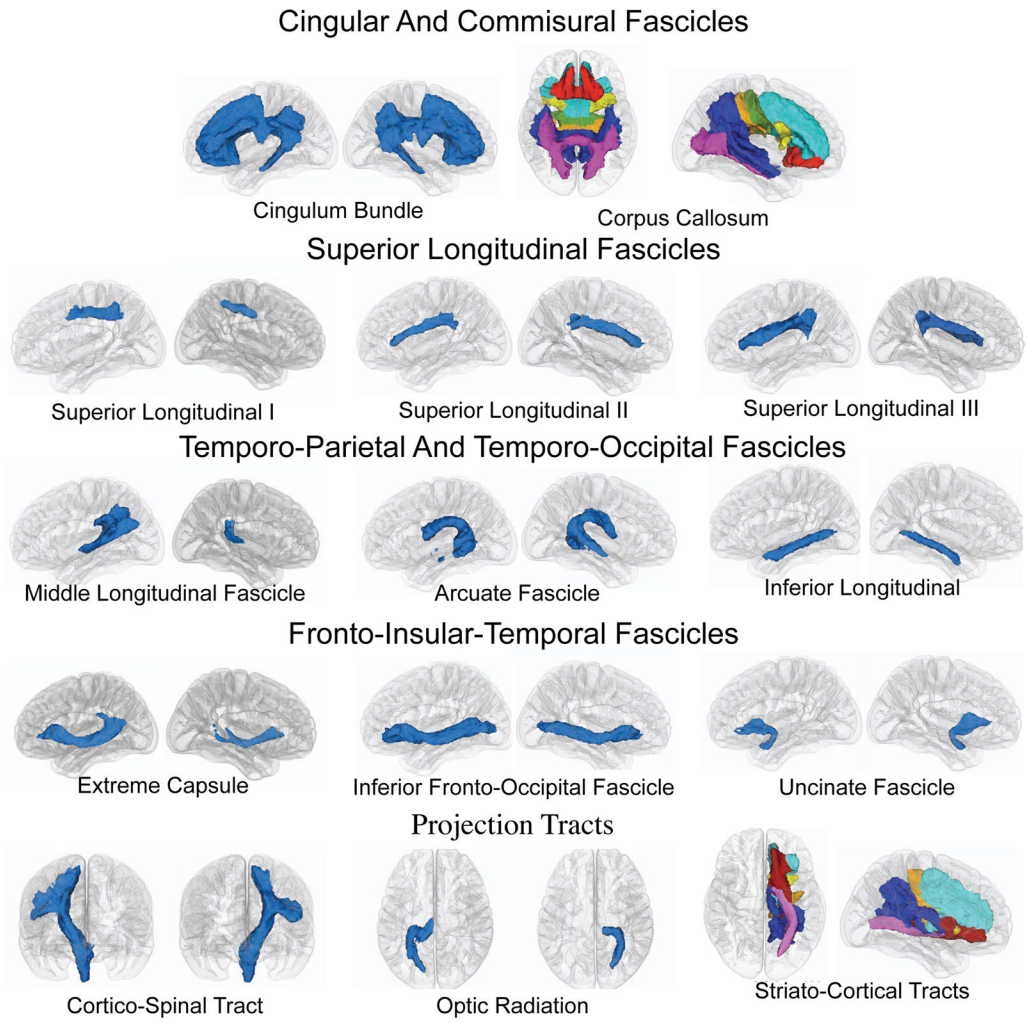

Fig. 2.

Iso-surfaces in blue (p-value = 0.01, corrected for multiple comparisons) shows 11 association, 8 projection and 7 commissural tracts.

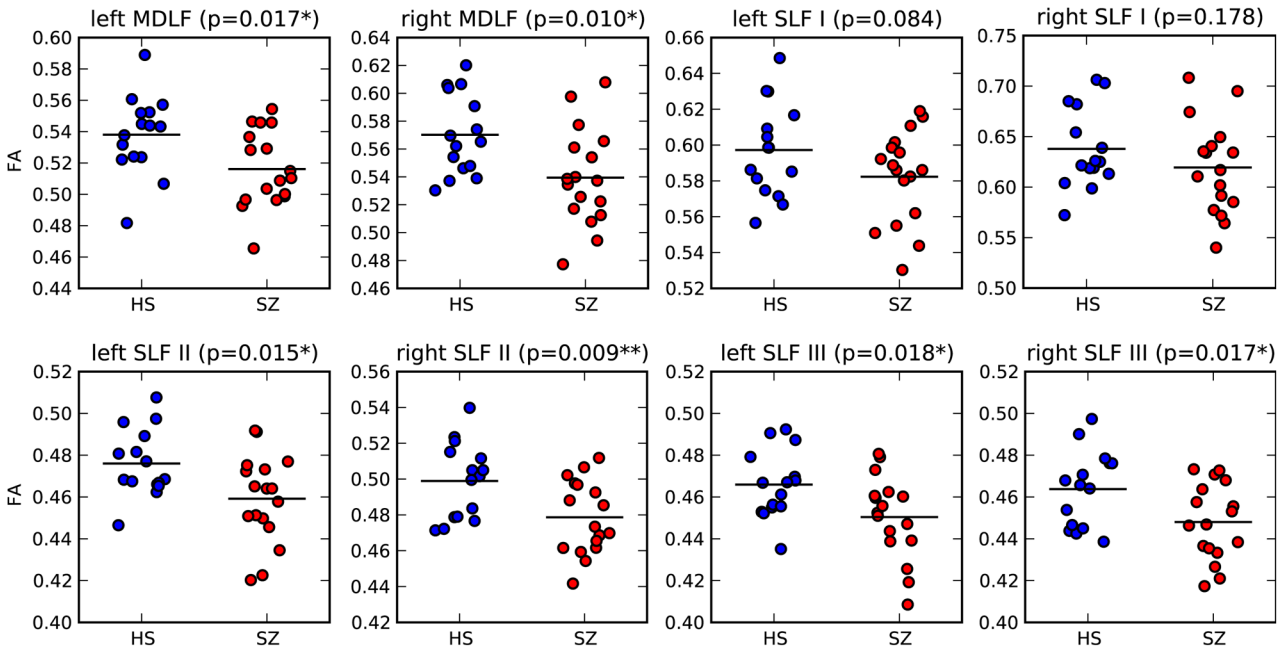

Fig. 3.

Compared mean FA over different tracts between healthy subjects (HS) and schizophrenic subjects (SZ). The p-value was computed using a t-test correcting for age. We observe significant differences on the left and right MdLF, and the left and right SLF II and III tracts. Notation: the * means p-value < .05 and ** p-value < .01

Table 1

Association Tract Definitions in WMQL: Cingulum bundle (CG); Extreme Capsule (EmC); and the fascicules: Superior Longitudinal (SLF) sections I to III; Arcuate (AF); Inferior occipito frontal (IOFF); Middle Longitudinal (MdLF); Uncinate (UF)

CB.side = only((cingular.**side or** cingular_cortex.**side**) **and** (middle_frontal.**side or** cuneus.**side or** entorhinal.**side or**superior_frontal.**side or** inferior_parietal.**side or** fusiform.**side or** medial_orbitofrontal.**side or** lateral_orbitofrontal.**side or** parahippocampal.**side or** precuneus.**side or** lingual.**side or** centrum_semiovale.**side**))
EmC.side = (**endpoints_in**(inferior_frontal.**side or** middle_frontal.**side**) **andendpoints_in**(inferior_parietal_lobule.**side**) **and**temporal.**side and** insula.**side**)**not in** hemisphere.**opposite**
SLF_I.side = (superior_parietal.**side and** precuneus.**side and** superior_frontal.**side**) **or**(superior_parietal.**side and** precuneus.**side and** superior_frontal.**side and** lateral_occipital.**side**) **not in** cingular.**side not in** temporal.**side not in** subcortical.**side not in** hemisphere.**opposite**
SLF_II.side =(inferior_parietal.**side or** supramarginal.**side or** lateral_occipital.**side**) **and endpoints_in**(middle_frontal.**side**)) **not in** temporal.**side not in** subcortical.**side not in** hemisphere.**opposite**
SLF_III.side = ((inferior_parietal.**side or** supramarginal.**side or** lateral_occipital.**side**) **and endpoints_in**(inferior_frontal.**side**)) **not in** temporal.**side not in** subcortical.**side not in** hemisphere.**opposite**
AF.side = (inferior_frontal.**side or** middle_frontal.**side or** precentral.**side**) **and** (superior_temporal.**side or** middle_temporal.**side**) **not in medial_of**(supramarginal.**side**)**not in** subcortical.**side not in** hemisphere.**opposite**
IOFF.side = (lateralorbitofrontal.**side and** occipital.**side**) **and** temporal.**side not in** subcortical.**side not in** cingular.**side not in** superior_parietal_lobule.**side not in** hemisphere.**opposite**
ILF.side = only(temporal.**side and** occipital.**side**) **and anterior_of**(hippocampus.**side**) **not in** parahippocampal.**side**
MdLF.side = only((temporal_pole.**side or** superior_temporal.**side**) **and** (inferior_parietal.**side or** superior_parietal.**side or** supramarginal.**side or** precuneus.**side or** (centrum_semiovale.**side and** superior_parietal.**side**) **or** (centrum_semiovale.**side and** inferior_parietal.**side**))
UF.side = insula.**side and** (inferior_frontal.**side or** middle_frontal.**side or** orbito_frontal.**side**) **and endpoints_in**(temporal.**side and anterior_of**(amygdala.**side**))