White matter changes in first episode psychosis and their relation to the size of sample studied: A DTI study

CITATION
1

10 AUTHORS, INCLUDING:

Tomas Melicher
University of Texas Health Science Center at Houston
8 PUBLICATIONS 2 CITATIONS
See Profile

Filip Spaniel
National Institute of Mental Health, Czech Republic
48 PUBLICATIONS 337 CITATIONS
See Profile

Jarda Tintera
Institute for Clinical and Experimental Medicine
29 PUBLICATIONS 493 CITATIONS
See Profile

Cyril Höschl
National Institute of Mental Health Klecany
211 PUBLICATIONS 1,883 CITATIONS
See Profile
**White matter changes in first episode psychosis and their relation to the size of sample studied: A DTI study**


**Abstract**

Background: White matter abnormality has been recently proposed as a pathophysiological feature of schizophrenia (SZ). However, most of the data available has been gathered from chronic patients, and was therefore possibly confounded by factors such as duration of the disease, and treatment received. The extent and localization of these changes is also not clear.

Methods: We examined a population of early stage SZ patients using diffusion tensor imaging (DTI). 77 SZ patients and 60 healthy controls (HCs) were included in the analysis using Tract-Based Spatial Statistics (TBSS). We have also analyzed 250 randomly created subsets of the original cohort, to investigate the relation between the result of TBSS analysis, and the size of the sample studied.

Results: We have found a significant decrease in fractional anisotropy (FA) in the patient group. This change is present in most major white matter (WM) tracts including the corpus callosum, superior and inferior longitudinal fasciculi, inferior fronto-occipital fasciculus, and posterior thalamic radiation. Furthermore, we identified a clear trend towards an increase in the number and spatial extent of significant voxels reported, with an increasing number of subjects included in the analysis.

Conclusion: Our study shows that FA is significantly decreased in patients at an early stage of schizophrenia, and that the extent of this finding is dependent on the size of studied sample; therefore underpowered studies might produce results with false spatial localization.

**1. Introduction**

Schizophrenia is a serious and disabling mental illness characterized by a wide array of symptoms and connected with structural and functional morphological brain changes (Wright et al., 2000; Ellison-Wright and Bullmore, 2008; Yan et al., 2012). The results from different brain imaging approaches indicate that a lower level of connectivity, either functional and/or anatomical, may be the essential neurobiological condition underlying the clinical symptoms of schizophrenia (Lynall et al., 2010; Samartzis et al., 2014).

Recent advances in magnetic resonance imaging (MRI) techniques have provided a useful tool for examining the pathological changes occurring in the brains of schizophrenia patients. The focus in schizophrenia research has recently shifted towards the investigation of white matter (WM) tracts connecting functionally distinct gray matter (GM) areas and the resulting defects in anatomical connectivity. Abnormal fiber myelination, and thus impaired connectivity, may account for numerous symptoms of schizophrenia, as stated in the disconnection hypothesis (Friston, 1998). In articles reviewing the research to date, authors report either a global WM volume decrease or diminished anisotropy in specific WM tracts and subtle decreases in myelin content (Fitzsimmons et al., 2013). They also implicate an abnormality in oligodendroglia as a possible underlying factor (Davis et al., 2003; Kunimatsu et al., 2012). The most widely replicated WM abnormalities are localized in the frontal lobes, the cingulum, and the corpus callosum (White et al., 2010). Two recent meta-analyses support these findings, showing a fractional anisotropy (FA) decrease in regions scattered across the brain (Yao et al., 2010; Bora et al., 2011).

The findings of WM abnormalities in schizophrenia are supported by a number of studies using either a standard voxel-based or region of interest (ROI) analysis of FA (Wright et al., 2000; Ellison-Wright and Bullmore, 2008; White et al., 2008; Takei et al., 2009; Sugranyes et al.,...
Additionally, we hypothesize that the previously reported and localized to isolated tracts. This would explain the observed inconsistency (White et al., 2010), the WM lesions are diffuse as opposed to being localized to isolated tracts. One of the essential questions that remains to be answered is whether these changes present in patients in their first psychotic episode. Regarding the fact that most of the studies performed to date have found WM impairment in chronically ill and extensively pharmacologically treated patients, one of the essential questions that remains to be answered is whether these changes present in medication-naive, first-episode schizophrenia patients. Our study addressed two fundamental questions: a) whether the WM changes are present in very early stages of schizophrenia spectrum disorders; and b) how extensive these changes are, and where are they localized in the brain.

The aim of this study was to determine the degree to which the integrity of WM tracts is compromised in the brains of first-episode psychosis patients. Firstly, we hypothesized that as a result of the proposed mechanisms of WM abnormalities in schizophrenia (such as immunological imbalance, genetic factors, or inflammation (Davis et al., 2003; White et al., 2010)), the WM lesions are diffused as opposed to being localized to isolated tracts. This would explain the observed inconsistency in the localization of WM defects reported in studies to date (Kanaan et al., 2005; Kyriakopoulos et al., 2008; Yao et al., 2010; Bora et al., 2011). Additionally, we hypothesized that the previously reported and commonly inconsistent associations of local WM changes with specific isolated tracts may be largely attributed to limited sample sizes, as the power may not be high enough to detect all of the changes. To illustrate this effect, we repeated our analysis on subsamples of the original dataset (sample size of 137 subjects), working with 250 random subsets of size ranging from 20 to 100 subjects in steps of 20 (50 random samples were generated for each size).

2. Methods

2.1. Participants

Patients included in this study are participating in the Early-Stage Schizophrenia Outcome study (ESO), a prospective trial conducted in a Prague and Central Bohemia surveillance area, investigating first-episode schizophrenia spectrum subjects. All of the 77 patients and 60 healthy volunteers that signed the informed consent and were administered with diffusion weighted MRI (DWI) were included in the Tract-Based Spatial Statistics (TBSS) analysis. Two experienced psychiatrists have diagnosed the patients as having schizophrenia (n = 19) or acute polymorphic psychotic disorder (n = 38) meeting the ICD-10 criteria. Patients with psychotic mood disorders (including schizoaffective disorder, bipolar disorder, and unipolar disorder with psychotic features) were not included in this study. First-episode psychosis (FEP) was defined by two criteria. For first contact patients, the duration of untreated psychosis (DUP) was ≤24 months. For previously treated patients, the total duration of illness was ≤18 months. There is no universally accepted definition of FEP; some authors suggest recent-onset psychosis as a more accurate term. Our criteria are in accordance with the consensual use of the term FEP, as laid out in a published review of the operational criteria (Breitborde et al., 2009). The mean duration of illness for the patient group was 5.5 months. All of the patients were receiving antipsychotic medication on the day of MRI scanning, and risperidone (n = 30) and olanzapine (n = 23), respectively, were the most common drugs used (Table 1). The mean daily chlorpromazine equivalent (Woods, 2003) of antipsychotics in the patient group was 306.9 mg (SD = 147.3). All of the patients underwent an initial evaluation including the Positive and Negative Syndrome Scale (PANSS) (Kay et al., 1987).

Healthy control subjects (HC) were recruited via local advertisements and they had a similar socio-demographic background as the FEP patients to whom they were matched for age and sex. The main exclusion criteria for control subjects were a personal lifetime history of any psychiatric disorder or substance abuse established by the Mini-International Neuropsychiatric Interview (M.I.N.I.) (Lecrubier et al., 1997), and any psychotic disorder in first or second-degree relatives.

Further exclusion criteria for both the patients and the healthy controls included current neurological disorders, a lifetime history of seizures or head injury with altered consciousness, intracranial hemorrhage, or neurological sequelae, a history of mental retardation, a history of substance dependence, and any contraindication for MRI scanning.

### Table 1

<table>
<thead>
<tr>
<th>Demographic and clinical characteristics of healthy controls and schizophrenia patients.</th>
<th>Controls (n = 60)</th>
<th>Patients (n = 77)</th>
<th>Test statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender — female (%)</td>
<td>32 (53%)</td>
<td>41 (56%)</td>
<td>x² = 0.38; df = 1; p = 0.53</td>
</tr>
<tr>
<td>Mean age, years (SD)</td>
<td>28.9 (7.1)</td>
<td>31.1 (7.7)</td>
<td>t-test; p = 0.07</td>
</tr>
<tr>
<td>Education, years (SD)</td>
<td>15.15 (2.5)</td>
<td>13.1 (2.4)</td>
<td>x² = 18.38; df = 2; p = 0.0008</td>
</tr>
<tr>
<td>Duration of illness, months (SD)</td>
<td>n/a</td>
<td>5.5 (5.9)</td>
<td></td>
</tr>
<tr>
<td>DUP, months (SD)</td>
<td>n/a</td>
<td>2.8 (3.9)</td>
<td></td>
</tr>
<tr>
<td>Antipsychotic treatment (%)</td>
<td>n/a</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>19.8 (9.5)</td>
<td>16.2 (6.9)</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>38.3 (14.2)</td>
<td>74.3 (26.5)</td>
<td></td>
</tr>
</tbody>
</table>

SD = standard deviation; DUP = duration of untreated psychosis; PANSS = Positive and Negative Syndrome Scale.

* Denotes a statistically significant difference.
of the subjects' aligned FA data were then projected onto this skeleton and the resulting data was fed into voxel-wise cross-subject statistics.

The FA maps were converted from NIfTI format to Analyze before performing analyses in SPM.

2.3. Statistical analyses

The demographic characteristics of the patient and control groups were compared using chi-square or Fisher's exact tests for categorical variables, and t-tests or Mann–Whitney U tests for continuous variables according to the distribution of the data. To investigate the potential confounding effect of age and years of education on FA, we performed analyses investigating their effect, and found no significant correlation.

For the WM skeleton creation we set the FA threshold to 0.15. Permutation methods using the Randomise tool (Nichols and Holmes, 2002; Winkler et al., 2014) were applied for statistical inference utilizing the Threshold free cluster enhancement (TFCE) (Smith and Nichols, 2009) analysis; family-wise error (FWE) correction of $q < 0.05$ was used to control for multiple testing (Nichols and Hayasaka, 2003).

To determine the influence of psychopathology, psychosis (and untreated psychosis) duration on FA, the normalized, segmented, and smoothed (with a 10 mm kernel) FA data were analyzed by SPM8 employing a general linear model (GLM). Due to the known correlation between age and FA (Pfefferbaum et al., 2000; Mori et al., 2007), also reflected in our dataset (Fig. 3), age was used as a nuisance covariate for all subsequent analyses. The influence of PANSS, its subscales (positive symptoms, negative symptoms, general psychopathology...
symptoms) and other covariates was treated by multiple regression analysis for either the whole group or for FEP patients and controls separately. Voxel-level inferences were used at $p \leq 0.005$ with family-wise error (FWE) correction, and clusters containing at least 50 contiguous voxels were considered as significant for all of the analyses.

2.4. Resampling

To determine the relationship between sample size and the extent of significant change in FA, we rerun the analysis on subsets of the original sample.

We used 5 different subsample sizes ($n = 20$ to $n = 100$ in steps of 20), with equal numbers of patients and HCs in each set. That is, we ran comparisons of 10 healthy controls vs 10 patients, 20 vs 20, 30 vs 30, 40 vs 40 and 50 vs 50. To gain an insight into the potential variability of the results due to variability of the random sample selection for each of these sample sizes, we generated $N_{\text{resample}} = 50$ different subsamples by randomly selecting the appropriate numbers of subjects from our main pool. The subsequent steps in each of these analyses were identical to the ones used for the main sample, i.e. we used TBSS, with the same parameters as in the primary analysis.

For the subsampled data we used voxel-wise uncorrected, cluster-wise FWE corrected, and TFCE multiple-comparison corrected statistical inference for a clearer illustration of the issue. For each type of analysis we report the mean number of significant voxels and the standard deviation of this mean (computed across the $N_{\text{resample}} = 50$ subsamples given the sample size $n$).

3. Results

We found a diffuse decrease in fractional anisotropy in patients with first episode psychosis compared to healthy controls (Figs. 1, 2). These changes are bilateral and generalized, affecting the genu, body, and splenium of corpus callosum, superior and inferior longitudinal fasciculi, inferior fronto-occipital fasciculus, posterior thalamic radiation, corona radiata and the white matter of frontal and parietal lobes (Table 2). Quantitatively, 14.5% of voxels within the WM skeleton mask showed a significant difference in the TFCE corrected analysis; and 13.7% in the

![Fig. 3. Negative association with age for the whole group of schizophrenia patients and control subjects, regions of negative association with age are displayed on a study specific T1-weighted template, on uncorrected $p \leq 0.001$. For three clusters, denoted by arrows, the associations survived FWE corrected voxel-threshold of $p \leq 0.05$.](http://dx.doi.org/10.1016/j.schres.2015.01.029)

Table 2
Clusters of voxels identified by the main TBSS analysis.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Voxels</th>
<th>$z$-Max $(1-p)$</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Location (JHU-ICBM-DTI-81 atlas)*</th>
<th>Location (visually)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>23,194</td>
<td>0.988</td>
<td>62</td>
<td>71</td>
<td>90</td>
<td>Posterior thalamic radiation</td>
<td>Body, splenium, genu of corpus callosum, SLF, ILF, inferior fronto-occipital fasciculus</td>
</tr>
<tr>
<td>2</td>
<td>214</td>
<td>0.954</td>
<td>42</td>
<td>76</td>
<td>96</td>
<td>None</td>
<td>Parietal WM</td>
</tr>
<tr>
<td>3</td>
<td>104</td>
<td>0.954</td>
<td>133</td>
<td>154</td>
<td>69</td>
<td>None</td>
<td>Frontal WM</td>
</tr>
<tr>
<td>4</td>
<td>68</td>
<td>0.952</td>
<td>58</td>
<td>66</td>
<td>115</td>
<td>None</td>
<td>Parietal WM</td>
</tr>
<tr>
<td>5</td>
<td>52</td>
<td>0.952</td>
<td>53</td>
<td>68</td>
<td>99</td>
<td>None</td>
<td>Parietal WM</td>
</tr>
<tr>
<td>6</td>
<td>28</td>
<td>0.95</td>
<td>53</td>
<td>69</td>
<td>106</td>
<td>None</td>
<td>Parietal WM</td>
</tr>
<tr>
<td>7</td>
<td>23</td>
<td>0.952</td>
<td>110</td>
<td>177</td>
<td>63</td>
<td>None</td>
<td>Frontal WM</td>
</tr>
<tr>
<td>8</td>
<td>20</td>
<td>0.952</td>
<td>122</td>
<td>72</td>
<td>113</td>
<td>None</td>
<td>Parietal WM</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>0.95</td>
<td>54</td>
<td>71</td>
<td>84</td>
<td>Posterior thalamic radiation</td>
<td>Posterior thalamic radiation</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>0.95</td>
<td>121</td>
<td>155</td>
<td>107</td>
<td>None</td>
<td>Frontal WM</td>
</tr>
</tbody>
</table>

$x, y, z$ – coordinates in MNI152 standard space; SLF = superior longitudinal fasciculus; ILF = inferior longitudinal fasciculus; and WM = white matter.

* Location of the “most significant” voxel.

Please cite this article as: Melicher, T., et al., White matter changes in first episode psychosis and their relation to the size of sample studied: A DTI study, Schizophr. Res. (2015), http://dx.doi.org/10.1016/j.schres.2015.01.029
Fig. 4. Results of the resampling analysis. Y-axis: number of voxels above the statistical threshold (group difference in FA between healthy controls and schizophrenic patients), as a function of subsample size. The number on x-axis for each bar indicates the total number of subjects (i.e. n = 20 for analysis comparing 10 random patients with 10 random HCs, etc.) The mean number of "significant" voxels is indicated by the bar height, the error bars show approximate 95% confidence interval (given by mean $\pm 2\times$SEM). From left to right the graphs correspond to the three approaches to statistical thresholding used: a, voxel-wise inference (p $b$ 0.05 uncorrected); b, TFCE (p $b$ 0.05 FWE corrected); c, cluster-wise inference (p $b$ 0.05 FWE corrected). Apart from the increase with sample size, note the higher variability in the case of the two latter thresholding approaches.
voxe-l-wise, uncorrected analysis. While this is still far from being significant for all WM voxels, visual inspection shows that the non-significant voxels typically lie in the minor tracts, where high inter-individual anatomical variability may complicate registration and lower the statistical power. However, we conjecture that for even larger sample sizes, these areas would show significant differences between the groups.

Regarding our hypothesis that the extent of change in FA between the two groups is dependent on the size of the sample studied, we repeatedly analyzed smaller, randomly created, subsets of our original sample. In these analyses, we found the results to be dependent on the cohort size (Table 3). There is a clear and strong trend toward an increasing number of significant voxels with the larger samples for both the uncorrected voxel-wise (Fig. 4a) and corrected TFCE analyses (Fig. 4b). Notably, the TFCE results show quite high variability across subsamples in terms of extent of above-threshold area, particularly for small sample sizes. For many subsamples no significant difference was detected, with 80, 66, 62, 56, and 24% subsamples giving no significant group-difference for sample sizes of 20, 40, 60, 80, and 100, respectively. We did not find any significant correlations between FA and either PANSS scores (positive, negative, general, or total) or DUP.

4. Discussion

Our results revealed two main findings. Firstly, we detected a widespread decrease in fractional anisotropy in most major WM tracts of patients with first episode psychosis. This finding supports the assumption of WM abnormalities at the very beginning of the clinical course of schizophrenia (Yao et al., 2013; Samartzis et al., 2014). The observed WM changes appear to be diffusely distributed throughout the brain, which indicates that the WM change is globally distributed and not localized to particular tracts. Secondly, the analysis of resampled randomly selected subsets of the whole sample supports our secondary assumption that the extent of detected change in FA is strongly dependent on the size of the studied sample. This finding strongly implies that studies with smaller numbers of subjects may produce results suggesting abnormalities falsely perceived as being regional. We identified a clear trend towards an increase in the number of significant voxels reported with an increase in the number of subjects analyzed. Particularly for the uncorrected voxel-wise analysis, it is reasonable to expect that the false positive rate is close to 5% independently of sample size, and to attribute the increase in the detected significant voxel count with larger samples to an increased true positive rate—i.e. correctly detecting larger a proportion of the affected area.

This explains why, even though the whole brain is affected, a lot of research performed to date on smaller subject samples has produced results localized to specific tracts. Further studies with higher numbers of patients analyzed, including multi-center studies and meta-analyses, are needed to address the question of size and localization of WM changes in the brains of patients with schizophrenia.

The pathophysiology underlying WM changes remains to be elucidated. However, our findings suggest that the WM deficit might be an endophenotype present in the very early stages of the disease, in drug-naïve patients (Liu et al., 2014), in a high risk population (Kubicki et al., 2013), and in healthy relatives of schizophrenia patients (Camchong, 2009; Wei et al., 2012). This is likely caused by a systemic process such as a genetic cause (Hakak et al., 2001; Karoutzou et al., 2007) (e.g. a polymorphism in one or more of the genes regulating myelination) or an immune system dysregulation (Müller and Schwarz, 2010) that might be brought about by an inflammatory process (Müller et al., 2012). It is important to stress that the FA changes reported in DTI studies may be confounded by various parameters including the cumulative dose of antipsychotics, duration of psychosis, DUP, or age of patients. Our findings are limited by the fact that they come from a cross-sectional study. Future studies should answer the question of the temporal dynamic of FA changes in schizophrenia patients and their relation to the duration of psychosis (and DUP), and antipsychotic treatment.

Our study has confirmed a global decrease in FA in schizophrenia patients in the early stages of the disease. This finding seems to be more pronounced in larger samples. However, because patients are generally not scanned upon admission and antipsychotic treatment is generally started as soon as possible it was not possible to obtain the MRI images from completely drug-naïve subjects. Thus, we are not able to rule out the effect of medication on FA. This limitation is offset by research indicating that antipsychotic medication actually increases FA in patients exposed to it (Reis Marques et al., 2014). We can therefore assume that the difference would be even more pronounced had the patient population not been exposed to antipsychotics. Another limitation of our study is the statistically significant difference between patients and HCs in years of education. The literature indicates no correlation between IQ and FA in schizophrenia patients (Kumra et al., 2004, 2005; Kuroki et al., 2006). Since the number of years of education is correlated with IQ, it can be considered a proxy measure (Matarazzo and Herman, 1984; Deary et al., 2007). We have verified this assumption for our sample by performing an analysis investigating the correlation between years of education and FA values, that found no significant correlation between these two variables.

Concerning the results of the analyses of subsamples utilizing TFCE with highly variable number of negative results, we attribute this high variability to the inherent dependence between TFCE image values of neighboring or even quite distant voxels (also observable for standard cluster-based thresholding, especially for lower thresholds), and believe this should be taken into consideration when interpreting TFCE analysis results; however a detailed analysis of this effect is outside the scope of this paper.

Role of funding source

The funding source had no role in study design; the collection, analysis, nor interpretation of data; writing of the paper; nor in the decision for publication.

Contributors

F. Spaniel, J. Horacek and T. Melicher designed the study, J. Tintera and I. Ibrahim acquired the imaging data, while F. Spaniel, T. Melicher, P. Mikolas, and P. Mohr participated in clinical examinations and scales. T. Novak analyzed the subject cohort demographics, T. Melicher (TBSS), J. Horacek (SPM), and J. Hlinka (resampling design) analyzed the data with help from P. Mikolas, who did programming and scripting, T. Melicher, J. Hlinka, and J. Horacek wrote the article, which all authors reviewed and approved for publication.

Conflict of interest

The authors report no conflicts of interest.

Acknowledgment

This study was supported by Charles University Grant Agency (GAUK) project No. 1352213, and by the Czech Science Foundation project No. 13-23940S.

Table 3

Table 3: Results of resampling analyses.

<table>
<thead>
<tr>
<th>n</th>
<th>Voxelwise</th>
<th>Clusters</th>
<th>TFCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>Mean</td>
<td>13198.8</td>
<td>15887.0</td>
<td>16607.3</td>
</tr>
<tr>
<td>S.D.</td>
<td>6635.8</td>
<td>5573.1</td>
<td>5044.8</td>
</tr>
</tbody>
</table>

There were 50 random samples analyzed for each method and sample size n (i.e. 750 analyses in total).

Here, we report the mean number of “significant” voxels for each sample size and inference method, and the corresponding standard deviation.