Neuroanatomical Abnormalities That Predate the Onset of Psychosis

A Multicenter Study

Andrea Mechelli, PhD; Anita Riecher-Rössler, MD; Eva M. Meisenzahl, MD; Stefania Tognin, BSc, MSc; Stephen J. Wood, PhD; Stefan J. Borgwardt, MD; Nikolaos Koutsouleris, MD; Alison R. Yung, MB, BS, PhD, FRANZCP; James M. Stone, BSc, MB, BS, MRCPsych, PhD; Lisa J. Phillips, MPsych, PhD; Patrick D. McGorry, MB, BS, MD, PhD, FRANZCP; Isabel Valli, MD; Dennis Velakoulis, MB, BS, FRANZCP; James Woolley, BSc, MBBS, MRCP, MRCPsych; Christos Pantelis, MB, BS, MD, MRCPsych, FRANZCP; Philip McGuire, BSc, MB, ChB, MD, PhD, FRCPsych

Context: People experiencing possible prodromal symptoms of psychosis have a very high risk of developing the disorder, but it is not possible to predict, on the basis of their presenting clinical features, which individuals will subsequently become psychotic. Recent neuroimaging studies suggest that there are volumetric differences between individuals at ultra-high risk (UHR) for psychosis who later develop psychotic disorder and those who do not. However, the samples examined to date have been small, and the findings have been inconsistent.

Objective: To assess brain structure in individuals at UHR for psychosis in a larger and more representative sample than in previous studies by combining magnetic resonance imaging data from 5 different scanning sites.

Design: Case-control study.

Setting: Multisite.

Participants: A total of 182 individuals at UHR and 167 healthy controls. Participants were observed clinically for a mean of 2 years. Forty-eight individuals (26.4%) in the UHR group developed psychosis and 134 did not.

Main Outcome Measures: Magnetic resonance images were acquired from each participant. Group differences in gray matter volume were examined using optimized voxel-based morphometry.

Results: The UHR group as a whole had less gray matter volume than did controls in the frontal regions bilaterally. The UHR subgroup who later developed psychosis had less gray matter volume in the left parahippocampal cortex than did the UHR subgroup who did not.

Conclusions: Individuals at high risk for psychosis show alterations in regional gray matter volume regardless of whether they subsequently develop the disorder. In the UHR population, reduced left parahippocampal volume was specifically associated with the later onset of psychosis. Alterations in this region may, thus, be crucial to the expression of illness. Identifying abnormalities that specifically predate the onset of psychosis informs the development of clinical investigations designed to predict which individuals at high risk will subsequently develop the disorder.
been reported, but the findings have been inconsistent; this may partly reflect the use of small samples. However, UHR individuals are hard to recruit, and it is difficult for any single center to scan a large sample. Sample size is a particular problem for the key comparison between UHR individuals who later develop psychosis and those who do not, which entails a further subdivision of the UHR sample according to clinical outcome. A potential solution is to conduct multicenter studies, with the pooling of data to produce a relatively large total sample. This approach has been successfully used in neuroimaging studies of other central nervous system disorders in which patient recruitment is difficult.21,22

In the present study, whole-brain MRIs were acquired from 5 MRI scanners in London, United Kingdom (2 sites); Basel, Switzerland; Munich, Germany; and Melbourne, Australia. The objective was to identify the most robust neuroanatomical abnormalities in individuals at UHR and to compare UHR participants who subsequently made a transition to psychosis with those who did not. At each site, participants were scanned at first clinical presentation and were observed clinically for a mean of 2 years so that they could then be subcategorized according to clinical outcome. The MRI data from each site were combined to form a large UHR sample, which was subdivided into individuals who had developed psychosis (UHR-T) and those who had not (UHR-NT). The MRI data from several matched healthy controls were also acquired at each site.

The first prediction, based on data from previous studies,11,15,18,23-27 was that the UHR group as a whole would show regional volumetric differences relative to controls that were qualitatively similar to those seen in patients with schizophrenia. We then tested the main hypothesis that UHR-T individuals would show differences in GMV relative to UHR-NT individuals in the inferior frontal, parahippocampal, and superior temporal cortices, the areas most frequently implicated in previous studies. Critically, these predictions were based on the results of previous single-center studies11,15,18,23-27 that used samples other than those reported in the present multisite investigation.

METHODS

SAMPLE

All the UHR individuals were recruited from specialized clinical services for people at high risk for psychosis. Individuals at UHR scanned at the 2 London sites (Institute of Psychiatry and Maudsley Hospital) were recruited via Outreach and Support in South London (OASIS). Individuals at UHR scanned in Basel were recruited through the clinic for early detection of psychosis (FEPSY) at the Psychiatric Outpatient Department, University Hospital. Those scanned in Munich were recruited through the Early Detection and Intervention Centre for Mental Crisis of the Department of Psychiatry and Psychotherapy, Ludwig-Maximilians-University. The UHR individuals scanned in Melbourne were recruited from the Personal Assessment and Crisis Evaluation (PACE) Clinic. Data were, thus, combined from 5 MRI scanners at the Institute of Psychiatry, the Maudsley Hospital, the University Hospital Basel, the Ludwig-Maximilians-University (Munich), and the Royal Melbourne Hospital. Some of the data from London, Basel, Munich, and Melbourne have been described separately in previous single-center studies.12-14,17,19,20,28-30 Data were from 182 patients at UHR.

All the UHR individuals at all the sites met the Personal Assessment and Crisis Evaluation criteria for UHR. Inclusion required the presence of 1 or more of the following features: (1) "attenuated" psychotic symptoms, (2) brief limited intermittent psychotic symptoms, or (3) a first-degree relative with a psychotic disorder, plus a marked decline in social or occupational functioning.27 None of the participants met the criteria for additional psychiatric disorders or learning disabilities. Details of the inclusion and exclusion criteria and clinical characteristics of the UHR individuals are reported in eTables 1, 2, 3, 4, 5, 6, and 7 (http://www.archgenpsychiatry.com). Most of the UHR group (168 of 182 individuals [92.3%]) had never taken antipsychotic agents or mood stabilizing drugs, and 14 (7.7%) had been exposed to antipsychotic agents (mean [SD] exposure time, 13.0 [19.3] months).

At each site, controls from the same geographic area as the UHR individuals were recruited through local advertisements. The control sample comprised 167 individuals and was comparable with the total UHR group for sex, age, and race/ethnicity. For all participants, the exclusion criteria were past or present diagnosis of psychiatric illness, previous treatment with antipsychotic drugs, medical illness, family history of psychiatric illness, past or present diseases of the central nervous system, alcohol or other substance abuse or dependence (defined using DSM-IV criteria), and pregnancy (eTable 1).

During a mean (SD) of 30.6 (10.4) months of follow-up subsequent to scanning, 48 of the UHR sample (26.4%) developed psychosis (UHR-T) and 134 did not (UHR-NT). Each site yielded a data set that included a UHR-T group, a UHR-NT group, and a control group, except 1 of the London data sets, from which the UHR-T group was too small (n=2) to be included in the combined UHR-T vs UHR-NT comparison (eTable 8).

MRI ACQUISITION

At all 5 sites, volumetric MRIs were acquired using a T1-weighted protocol. At 4 sites, the scanner field strength was 1.5 T, and at 1 site it was 3 T. Three sites used General Electric scanners (Milwaukee, Wisconsin), and 2 used Siemens scanners (Erlangen, Germany). The details of the MRI acquisition sequence varied among scanners (eTable 9).

DATA ANALYSIS

Sociodemographic and Clinical Parameters

Differences in demographics and clinical profile between groups were examined using 1-way analysis of variance for parametric data and a χ² test for nonparametric data using a commercially available software program (SPSS, version 17.0 for Windows; SPSS Inc, Chicago, Illinois) (Table 1).

Preprocessing

We examined group-related differences in GMV using voxel-based morphometry, as implemented in SPM8 software (http://www.fil.ion.ucl.ac.uk/spm) running under MATLAB 7.1 (The MathWorks, Inc, Natick, Massachusetts). First, T1-weighted volumetric images were preprocessed using the DARTEL (diffeomorphic anatomical registration using exponentiated lie algebra)33 SPM8 toolbox. This approach involves the creation of a study-specific template and the segmentation of each in-
individual image using such a template, with the aim of maximizing accuracy and sensitivity. The following steps were followed for voxel-based morphometry preprocessing: (1) checking for scanner artifacts and gross anatomical abnormalities for each subject, (2) setting the image origin to the anterior commissure, (3) using the DARTEL toolbox to produce a high-dimensional normalization protocol, (4) checking for homogeneity across the sample, and (5) using standard smoothing (ie, 8 mm). A “modulation step” was also included in the normalization to preserve the information about the absolute gray matter values. After this preprocessing, we obtained smoothed, modulated, normalized data that were used for the statistical analysis.

STATISTICAL ANALYSIS

We performed 2 statistical analyses using SPM8 software. First, an analysis of variance was used to compare gray matter images from UHR individuals (UHR-T and UHR-NT combined) and controls. In this analysis, scanner site was modeled as a factor, resulting in 10 experimental groups. Second, we performed an analysis of variance to compare gray matter images from UHR individuals who later became psychotic (UHR-T), UHR individuals who did not become psychotic (UHR-NT), and controls. In this analysis, scanner site was again modeled as a factor, resulting in 14 experimental groups (1 UHR-T group was too small for analysis [n = 2]; compare with the UHR-NT group [n = 19]; see previously herein). Including scanner site as a factor in the statistical analysis allowed us to (1) model scanner-related variance in the data, which had the effect of reducing error variance and increasing statistical sensitivity, and (2) examine the impact of scanner site by testing for scanner effects and scanner × group interactions. To assess how much of the interindividual variance in regions that differed between UHR-T and UHR-NT was explained by diagnostic group and scanner site, respectively, we used the η² measure of effect size in SPSS. In both analyses, we modeled age, sex, race/ethnicity, and use of medication as covariates of no interest to reduce the potential impact of these variables on the findings. To identify regionally specific changes that were not confounded by global differences, we used the proportional scaling option. Statistical inferences were made at P < .05 after familywise error (FWE) correction, with an extent threshold of 5 voxels.

REGION-OF-INTEREST ANALYSES

In addition to the whole-brain analysis, regions of interest (ROIs) were used to examine between-group differences in areas in which volumetric abnormalities have previously been identified in studies of people at high risk for psychosis or patients with first-episode psychosis. These ROIs were the left parahippocampal gyrus (MNI [Montreal Neurological Institute] coordinates x, y, z, −23, 6, and −20), the right inferior frontal gyrus (45, 37, and 0), the left superior temporal gyrus (−49, −31, and 2), and the left superior temporal gyrus (−49, −31, and 2). We used the coordinates from the previous studies that had reported the most significant effects in these regions; none of these previous studies had included data from individuals who participated in the present investigation. Using the SimpleROIBuilder toolbox (http://www.fil.ion.ucl.ac.uk/spm/ext/), we created a mask that included the 3 chosen ROIs. This mask consisted of 3 spheres with a radius of 8 mm corresponding to the 3 ROIs and a total of 758.71 voxels. Within the mask, statistical inferences were made at P < .05 after FWE correction for multiple comparisons.

SOCIODEMOGRAPHIC AND CLINICAL PARAMETERS

No statistically significant differences were noted among the UHR-T, UHR-NT, and control groups in age, sex, total GMV, and race/ethnicity (Table 1).

DIFFERENCES BETWEEN THE UHR AND CONTROL GROUPS

The UHR group had less GMV than did controls (at P < .05 after FWE correction) in 3 areas of the frontal cortex: the medial orbital gyrus and the gyrus rectus bilaterally and the right anterior cingulate gyrus (Figure 1 and Table 2). In these regions, no significant effect of medication was noted even at trend level (P < .05 uncorrected). There were no areas in which UHR individuals had more GMV than did controls.

Table 1. Demographic Data and Global Brain Volumes of Study Samples

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Healthy Controls (n = 167)</th>
<th>UHR-T Group (n = 48)</th>
<th>UHR-NT Group (n = 134)</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD), y</td>
<td>23.5 (4.2)</td>
<td>22.7 (4.5)</td>
<td>23.3 (5.3)</td>
<td></td>
</tr>
<tr>
<td>Sex, No.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>104</td>
<td>15</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>63</td>
<td>33</td>
<td>83</td>
<td></td>
</tr>
<tr>
<td>Race/ethnicity, No.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>140</td>
<td>42</td>
<td>106</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>11</td>
<td>1</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>10</td>
<td>0</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Mixed</td>
<td>6</td>
<td>5</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Handedness, No.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right</td>
<td>147</td>
<td>44</td>
<td>126</td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>14</td>
<td>3</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Ambidextrous</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMV, mean (SD), mm³</td>
<td>0.948 (0.11)</td>
<td>0.945 (0.08)</td>
<td>0.937 (0.10)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: GMV, gray matter volume; UHR-NT, ultra-high risk without disease transition; UHR-T, ultra-high risk with disease transition.
DIFFERENCES BETWEEN UHR INDIVIDUALS WHO DID AND DID NOT DEVELOP PSYCHOSIS

Analysis by ROI revealed reduced gray matter volume in the UHR-T group relative to the UHR-NT group in the anterior part of the left parahippocampal gyrus (bordering the uncus) (MNI coordinates x = −21, y = 6, and z = −27; P = .03; z = 3.35; and cluster size = 6 voxels) (at P < .05 after FWE correction) (Figure 2), where the difference between UHR-T and UHR-NT accounted for 14% of the total variance. In this region, no significant effect of medication was noted even at trend level (P < .05 uncorrected). Plotting of gray matter values revealed that this reduction was evident in each of the 4 sites examined for this contrast (Figure 2). There were no significant differences in the other ROIs (ie, in the right inferior frontal gyrus and the left superior temporal gyrus).

To examine the predictive value of GMV in the left parahippocampal gyrus, we extracted gray matter values from the peak voxel and performed a series of cross-validation analyses using a predictive linear model in SPSS software. This involved developing a predictive model based on a data set from a single scanner site and testing it in each of the other data sets; the average predictive accuracy was 62% (sensitivity = 61% and specificity = 69%). We also performed a 3-fold cross-validation analysis irrespective of scanner site that involved developing a predictive model based on two-thirds of the total sample and testing it in the remaining one-third; this yielded an average accuracy of 67% (sensitivity = 68% and specificity = 66%).

We used MRI to study a large sample of UHR individuals created by pooling data from 5 sites. The UHR individuals were observed clinically subsequent to scanning and were subcategorized according to which individuals developed psychosis and which did not.

On the basis of previous MRI studies of smaller UHR samples collected at single sites,11,15,18,23-27 we first tested the hypothesis that the UHR group as a whole would show volumetric abnormalities relative to controls that were qualitatively similar to those seen in patients with schizophrenia. Consistent with this prediction, the UHR group expressed significant reductions in GMV in the prefrontal and anterior cingulate cortices (P < .05 after FWE correction), areas that have been consistently implicated in volumetric neuroimaging studies of schizophrenia.37 In contrast, we did not identify areas where there was more GMV in the UHR sample than in controls.

The main prediction was that UHR individuals who later developed psychosis would show differences in regional GMV in the inferior frontal, parahippocampal, and superior temporal cortices compared with those who did not become psychotic. This hypothesis was, in part, confirmed: the UHR-T subgroup showed relatively reduced GMV in the anterior part of the left parahippocampal gyrus, bordering the hippocampal uncus (P < .05 after FWE correction). In this area, there was less gray matter in UHR-T individuals than in controls (P < .05 corrected) but no significant difference between the UHR-NT and

Table 2. MNI Coordinates and z Scores for Regions Showing Differences in Gray Matter Volume Between the UHR and Healthy Control Group

<table>
<thead>
<tr>
<th>Healthy Control &gt; UHR Area</th>
<th>Hemisphere</th>
<th>MNI Coordinates</th>
<th>Cluster Size, No. of Voxels</th>
<th>z Score</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medial orbital gyrus</td>
<td>Right</td>
<td>7 40.5 −16.5</td>
<td>130</td>
<td>5.44</td>
<td>.001</td>
</tr>
<tr>
<td>Cingulate gyrus</td>
<td>Right</td>
<td>12 58.5 9</td>
<td>50</td>
<td>5.35</td>
<td>.001</td>
</tr>
<tr>
<td>Medial orbital gyrus</td>
<td>Left</td>
<td>−18 52.5 −16.5</td>
<td>30</td>
<td>4.94</td>
<td>.005</td>
</tr>
<tr>
<td>Gyrus rectus</td>
<td>Right</td>
<td>3 25.5 −25.5</td>
<td>73</td>
<td>4.94</td>
<td>.01</td>
</tr>
<tr>
<td>Gyrus rectus</td>
<td>Left</td>
<td>−4 25.5 −22.5</td>
<td>4.80</td>
<td>.02</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: MNI, Montreal Neurological Institute; UHR, ultra-high-risk.

a P < .05 after familywise error correction.

b No clusters were detected for the UHR > control contrast.

Figure 1. Differences between the ultra-high-risk (UHR) and control (HC) groups. The images show the medial orbital region, where the total UHR sample showed reduced gray matter volume relative to HCs (P < .05 after familywise error correction).
control groups (eTable 10). A series of cross-validation analyses also revealed that GMV in this region allowed discrimination between UHR-T and UHR-NT with an accuracy of up to 67%. In contrast to some previous single-center studies,16,36 we did not find statistically significant differences in the inferior frontal and superior temporal gyri.

Reductions in parahippocampal volume have been reported in high-risk individuals relative to controls,36 as have alterations in parahippocampal function.38 Over time, reductions in parahippocampal volume have been described in high-risk individuals with transient or isolated psychotic symptoms,16 and longitudinal reductions have been described in high-risk individuals who developed psychosis.15 Moreover, cross-sectional comparisons indicate that patients with first-episode psychosis have smaller parahippocampal volumes than do controls and UHR individuals.26 Meta-analyses suggest that the parahippocampal region is one of the most robust sites of volume reduction in chronic schizophrenia.37,39-41 Contemporary animal models of psychosis propose that altered parahippocampal activity drives subcortical dopamine dysfunction.42 These observations are consistent with the notion that the parahippocampal cortex is critically implicated in psychotic disorders.

At the time of scanning, the UHR-T and UHR-NT groups were clinically indistinguishable; the volumetric differences observed between them could be interpreted as neurobiological markers of an especially increased vulnerability to psychosis or as early manifestations of a neuropathologic process underlying the transition to psychosis. Because the data in the present study were collected at a single cross-sectional time point, we cannot determine at what stage these differences first emerged. This issue could be addressed by longitudinal neuroimaging studies of individuals at different time points in the prodromal phase of psychosis.

A limitation of this study is that the data were collected using different scanners and different acquisition sequences.22,43,44 Nevertheless, we are confident that the results do not represent an artifact due to the use of different scanners for several reasons. First, we sought to control for these effects by using only MRI data collected with T1-weighted sequences and by modeling scanner site as an independent factor in the statistical analysis. Second, a comparable proportion of controls, nonconverters, and converters was scanned at each site, except for 1 of the London data sets, which was, therefore, excluded from the UHR-T vs UHR-NT comparison. Third, plotting of gray matter values suggested that the differences between the UHR-T and UHR-NT groups were evident not only across different centers but also within each site and, therefore, cannot be explained by interscanner differences (Figure 2). Fourth, when we examined the impact of scanner site, we found no evidence of either scanner effects or scanner × group interactions in regions that differed between groups, even when lowering the statistical threshold to $P < .05$ (uncorrected). Fifth, this approach to the integration of multiscanner data in the same statistical model has been used successfully in previous studies$^{22,43-45}$ that also combined different scanners and acquisition sequences; these studies typically found that scanner differences were substantially less than group differences. Consistent with this,
we found that scanner-related variance (11%) was less than group-related variance (14%) in the left parahippocampal region, which differed between the UHR-T and UHR-NT groups. The impact of the use of different scanners and acquisition sequences could be further controlled for by scanning the same individuals at each site, but this was not feasible in the present investigation.

Data from the present study suggest that in the future it may be possible to use MRI to facilitate the prediction of psychosis in those at high risk.46 This would be particularly useful in the clinical management of individuals at UHR because it is difficult to predict which individuals will go on to develop psychosis on the basis of their clinical features.4 As a result, it is not possible to focus the provision of clinical resources, such as putatively preventive treatments,77 to the subgroup of UHR individuals who will later become psychotic. Clinical application of neuroimaging in this context requires the identification of predictive markers at an individual level, whereas the present data represent group differences. Image analysis methods that classify individuals according to patterns of data associated with diagnostic categories may provide a means of addressing this issue.46

In conclusion, UHR is associated with alterations in regional GMV, and in this population, reductions in the parahippocampal region may be specifically linked to the later onset of psychosis. These findings suggest that neuroimaging data may facilitate the prediction of illness in individuals at high risk for psychosis and may inform the development of new interventions designed to delay or prevent its onset.

Submitted for Publication: June 11, 2010; final revision received November 4, 2010; accepted December 9, 2010.

Author Affiliations: Department of Psychology Studies, Institute of Psychiatry, King’s College London, London, United Kingdom (Drs Mechelli, Stone, Valli, and McGuire, Ms Tognin, and Mr Woolley); University Psychiatry Outpatient Department, Psychiatric University Clinics, University Hospital Basel, Basel, Switzerland (Drs Riecher-Rossler and Borgwardt); Departments of Psychiatry and Psychotherapy, Ludwig-Maximilians-University, Munich, Germany (Drs Meisenzahl and Koutsouleris); Melbourne Neuropsychiatry Centre, Department of Psychiatry, The University of Melbourne and Melbourne Health, Melbourne, Victoria, Australia (Drs Wood and Pantelis and Mr Velakouls); School of Psychology, University of Birmingham, Edgbaston, Birmingham, United Kingdom (Dr Wood); Orygen Research Centre (Drs Yung and McGorry) and Psychological Sciences (Dr Phillips), University of Melbourne; and Centre for Mental Health, Imperial College London, Hammersmith Hospital, London (Dr Stone).

Correspondence: Stefania Tognin, BSc, MSc, Department of Psychology Studies, PO Box 67, Institute of Psychiatry, King’s College London, De Crespigny Park, London SE5 8AF United Kingdom (stefania.tognin@kcl.ac.uk). Financial Disclosure: None reported.

Funding/Support: This study was supported by a National Alliance for Research on Schizophrenia and Depression Independent Investigator Award (Dr Mechelli).


Additional Contributions: The FEPSY research group in Basel assisted with data collection.