# Supplementary Methods

## MRI Data Preprocessing

The 3D sMRI data were preprocessed using CAT12 toolbox version 1364 (<http://www.neuro.uni-jena.de/cat/>) within SPM12 software (<http://www.fil.ion.ucl.ac.uk/spm>) on the MATLAB R2016b platform (<https://www.mathworks.com/products/matlab.html>). The preprocessing steps included: First, the sMRI data were corrected for the bias-field inhomogeneities. Second, sMRI data were segmented into gray matter, white matter, and cerebral spinal fluid. Third, the segmented images were further spatially normalized into the Montreal Neurological Institute (MNI) space using the DARTEL algorithm [1]. Fourth, the normalized gray matter images were multiplied by the Jacobian determinant to calculate the absolute gray matter volume (GMV) and were finally resampled to a voxel size of 1.5mm × 1.5mm × 1.5mm. The total intracranial volume (TIV) of each subject was also measured. Finally, a Gaussian kernel with 8-mm full-width at half-maximum (FWHM) was applied to smooth the GMV images spatially.

## Data Harmonization and Feature Extraction

Combat harmonization was performed to eliminate the systematic bias in the GMV maps (both for unsmoothed and smoothed GMV) and PANSS scores (including 7 positive items, 7 negative items, and 14 general cognitive items) across four sites. Combat harmonization was originally developed for gene expression microarray data analysis and recently has been reformulated to harmonize neuroimaging data across different sites [2-4]. Combat model supposed that the expected value of the features can be modeled as a linear combination of the site effects and biological effects, whose error term was mediated by additional site-specific scaling factors [3]. Combat harmonization employed the Bayesian regression method to discover and correct the system differences among multivariate data collected from multiple sites in our study and can simultaneously remove additive and multiplicative bias across the sites while preserving the biological variation of interest. The formula of the Combat model showed as follows:

$y\_{inv}=α\_{v}+X\_{in}β\_{v}+γ\_{iv}+δ\_{iv}ε\_{inv}$ [1]

where $y\_{inv}$ is the $ v^{th}$ GMV (PANSS items scores) of the $n^{th}$ individual in the $i^{th}$ site, $a\_{ν} $is the overall mean of the$ v^{th} $GMV (PANSS items scores),$X\_{in}$ are the covariates of interest of the $n^{th}$ individual in the$ i^{th}$ site, and $β\_{v} $is the effect of the covariates of interest of the $v^{th} $GMV (PANSS items scores). Additionally, $γ\_{iv} $and $δ\_{iv}$ indicate the addition and multiplication effects of $ⅈ^{th}$site for the $v^{th} $GMV (PANSS items scores).

After the combat harmonization, to further remove the effects of confounders, the age, gender, and TIV were regressed for the GMV data (both for unsmoothed and smoothed GMV), and the age and gender were regressed for PANSS items scores data.

Finally, a principal component analysis (PCA) was carried out for the unsmoothed GMV images, and the top 464 principal components that explain 95% of the variance were chosen for the following clustering analysis.

## Subtype Identification

The GMV (464 features) and PANSS data (30 features) of the recruited schizophrenia patients further underwent clustering analysis using K-means and Hierarchical clustering algorithms, respectively. For each subtyping, a grid search strategy was used to determine the hyperparameters for K-means and Hierarchical algorithms. The hyperparameters for K-means included: distance metric ('sqeuclidean', 'correlation', 'cityblock', 'cosine'), method for choosing initial cluster centroid positions ( 'cluster', 'uniform', 'plus'), and number of times to repeat clustering using new initial cluster centroid positions (1000). The hyperparameters for Hierarchical clustering included: distance metric ('euclidean', 'squaredeuclidean', 'seuclidean', 'minkowski', 'chebychev', 'jaccard', 'spearman', 'correlation', 'cityblock', 'cosine'), Algorithm for computing distance between clusters ('average', 'centroid', 'complete', 'median', 'single', 'ward', 'weighted'). The common hyperparameters for both K-means and Hierarchical clustering included: the number of subtypes (from 2-10).

A 10 randomization 5-fold cross-validation strategy was used to estimate and control sample selection bias during clustering. Specifically, in each of the 10 random samplings, we randomly split the patients' data into 5 folds, in which 4 folds were used for each clustering, resulting in 10×5 clustering results for each subtyping task. Then we ensembled the 50 clustering results into one subtyping according to label probability across the 50 shuffles with the following steps: first, because the label (subtype) indices were arbitrarily assigned in either K-means or hierarchical clustering, we defined the 1st clustering labels as the reference, then we iteratively reassigned the subtype indices of the target shuffle by maximum the Jaccard similarity coefficient between the reference and the target clustering labels with the following equation [1].

$J=\frac{\sum\_{i=1}^{k}A\_{i}∩B\_{i}}{A∪B}$ [1]

J represents the Jaccard similarity score, k represents the maximum label index, A and B represent the reference and target samples, and Ai and Bi represent the reference and target samples with label index i, respectively.

Second, after label reassignment, we calculated the probability of each label (subtyping probability) across the 50 shuffles for each subject, and the label with the maximum probability was finally assigned to this subject.

Third, to determine the best hyperparameters (distance, number of subtypes, initial centroid, etc.) for each subtyping model, we introduced two commonly used criteria in clustering analysis: Calinski-Harabasz Index (CHI) and Adjusted Rand Index (ARI). CHI evaluates the ratio of between-cluster variance and within-cluster variance (clustering distinguishability), and well-defined clusters are expected to have a large between-cluster variance and a small within-cluster variance. However, the CHI can not evaluate the repeatability (stability) of clusterings between different samplings, which ARI can realize. ARI computes the normalized similarity measure between predicted and true clusterings or between two predicted clusterings with overlapped samplings. We calculated the average CHI and ARI across 50 shuffles for each model. To account for both clustering distinguishability and repeatability, we use the multiplication of the average CHI and ARI (CHI\*ARI) to determine the best clustering hyperparameters (highest CHI\*ARI) for each subtyping model. The labels of the best model were used to assign the subtype of each subject.

The clustering analyses were performed using the house-coded scripts developed on MATLAB 2016b.

## Statistical Analysis

One-way analysis of variance (ANOVA) was performed to explore the GMV (smoothed GMV after ComBat harmonization and confounders regression) differences among the schizophrenia subtypes and HCs in the total cohort (voxel-wise family-wise error [FWE] corrected, *P* < 0.05), and in each site (voxel-wise *P* < 0.001, cluster-wise FWE corrected *P*< 0.05), respectively. The Mann-Whitney U test was used to explore the differences in PANSS items scores between the schizophrenia subtypes in the total cohort and each site (*P*<0.05, Bonferroni corrected), respectively.

A Kappa test was employed to explore the consistency of subtyping results between different subtyping features (GMV vs. PANSS) or between different subtyping methods (K-means vs. Hierarchical) (*P* < 0.05, Bonferroni corrected).

To explore the site effects on the between-subtype difference in GMV and PANSS, the T-distribution (two-sample t-test) of the between-subtype difference across voxels in GMV and Z-distribution (Mann-Whitney U test) across PANSS items were extracted. Then a Spearman correlation was used to test the association in T-distribution (or Z-distribution ) between each pair of the four sites and between each site and the total cohort (*P* < 0.05, Bonferroni corrected).

Kolmogorov–Smirnov test was used to explore the normal distribution of demographic data. One-way ANOVA (if normal distribution) or Kruskal-Wallis test (otherwise) was performed to explore the differences in age and TIV between schizophrenia subtypes and HCs (*P* < 0.05). Mann-Whitney U test was used to explore the differences in clinical measures between schizophrenia subtypes.

**References for Supplementary Methods**

[1] Ashburner J. A fast diffeomorphic image registration algorithm. Neuroimage. 2007;38:95-113.

[2] Fortin JP, Parker D, Tunc B, Watanabe T, Elliott MA, Ruparel K, et al. Harmonization of multi-site diffusion tensor imaging data. Neuroimage. 2017;161:149-70.

[3] Fortin JP, Cullen N, Sheline YI, Taylor WD, Aselcioglu I, Cook PA, et al. Harmonization of cortical thickness measurements across scanners and sites. Neuroimage. 2018;167:104-20.

[4] Radua J, Vieta E, Shinohara R, Kochunov P, Quide Y, Green MJ, et al. Increased power by harmonizing structural MRI site differences with the ComBat batch adjustment method in ENIGMA. Neuroimage. 2020;218:116956.

# Supplementary Figures

****

**Supplementary Fig. 1: Subtypes consistency between different features and between different clustering methods in the BrainGluSchi site.** A Kappa test was employed to explore the consistency of subtypes using different subtyping features (GMV vs. PANSS) or between different subtyping methods (K-means vs. Hierarchical) in the BrainGluSchi site. **(A)** GMV vs. PANSS subtypes consistency using Kmeans clustering; **(B)** GMV vs. PANSS subtypes consistency using Hierarchical clustering; **(C)** GMV subtypes consistency between Kmeans and Hierarchical clustering methods; **(D)** PANSS subtypes consistency between Kmeans and Hierarchical clustering methods.



**Supplementary Fig. 2: Subtypes consistency between different features and between different clustering methods in the COBRE site.** A Kappa test was employed to explore the consistency of subtypes using different subtyping features (GMV vs. PANSS) or between different subtyping methods (K-means vs. Hierarchical) in the COBRE site. **(A)** GMV vs. PANSS subtypes consistency using Kmeans clustering; **(B)** GMV vs. PANSS subtypes consistency using Hierarchical clustering; **(C)** GMV subtypes consistency between Kmeans and Hierarchical clustering methods; **(D)** PANSS subtypes consistency between Kmeans and Hierarchical clustering methods.



**Supplementary Fig. 3: Subtypes consistency between different features and between different clustering methods in the TIANJIN site.** A Kappa test was employed to explore the consistency of subtypes using different subtyping features (GMV vs. PANSS) or between different subtyping methods (K-means vs. Hierarchical) in the TIANJIN site. **(A)** GMV vs. PANSS subtypes consistency using Kmeans clustering; **(B)** GMV vs. PANSS subtypes consistency using Hierarchical clustering; **(C)** GMV subtypes consistency between Kmeans and Hierarchical clustering methods; **(D)** PANSS subtypes consistency between Kmeans and Hierarchical clustering methods.

****

**Supplementary Fig. 4: Subtypes consistency between different features and between different clustering methods in the GUANGZHOU site.** A Kappa test was employed to explore the consistency of subtypes using different subtyping features (GMV vs. PANSS) or between different subtyping methods (K-means vs. Hierarchical) in the GUANGZHOU site. **(A)** GMV vs. PANSS subtypes consistency using Kmeans clustering; **(B)** GMV vs. PANSS subtypes consistency using Hierarchical clustering; **(C)** GMV subtypes consistency between Kmeans and Hierarchical clustering methods; **(D)** PANSS subtypes consistency between Kmeans and Hierarchical clustering methods.

# Supplementary Tables

**Supplementary Table 1. The demographic and clinical characteristics of the subjects clustered using K-means method based on GMV data.**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Sites  |  | Subtype 1 | Subtype 2 | HCs | F/Z/χ2 value | P value |
| BrainGluschi | Number  | 40 | 31 | 44 |  | - |
| Age(year) | 35.60±12.75 | 32.52±12.26 | 34.32±10.58 | 0.594 | 0.554 |
|  | Male/Female | 35/5 | 28/3 | 38/6 | 0.306 | 0.937 |
|  | TIV(ml) | 1514.09±138.83 | 1519.04±153.90 | 1527.29±117.85 | 0.102 | 0.903 |
|  | **Total CPZ** | **212 (75, 400)** | **323 (200, 600)** | **-** | **-0.2768** | **0.006\*** |
|  | Age when SZ was firstly diagnosed | 19 (17, 24)  | 20 (18, 24) | - | -0.931 | 0.352 |
|  | Age when symptoms firstly appeared | 18 (16, 24) | 20 (17, 24) | - | -1.144 | 0.252 |
|  | Age when therapy was firstly received | 19 (17, 25) | 21 (16, 25) | - | -0.815 | 0.415 |
|  | Ill duration | 11 (2, 23) | 7 (3, 14) | - | -1.374 | 0.170 |
|  | Education level  | 3 (3, 4) | 3 (2, 4) | - | -0.930 | 0.352 |
| COBRE | Number  | 50 | 30 | 78 |  | - |
| Age | 36.28±11.77 | 38.53±13.74 | 37.12±10.85 | 0.346 | 0.708 |
|  | Male/Female | 42/8 | 22/8 | 56/22 | 2.624 | 0.274 |
|  | TIV | 1513.98±138.28 | 1465.40±142.66 | 1497.81±159.53 | 0.987 | 0.375 |
|  | Total CPZ | 300 (187, 557.5) | 500 (145, 640) | - | -1.218 | 0.223 |
|  | Age when SZ was firstly diagnosed | 19 (17, 25) | 20 (17, 23.5) | - | -0.503 | 0.615 |
|  | Age when symptoms firstly appeared | 18 (15.75, 24) | 20 (17, 23.5) | - | -1.676 | 0.094 |
|  | Age when therapy was firstly received | 19 (17, 25) | 22 (19, 31) | - | -1.748 | 0.080 |
|  | Ill duration | 13.5 (6, 20.25) | 10 (5.75, 25.5) | - | -0.587 | 0.557 |
|  | Education level  | 3 (3, 4) | 4 (2.5, 4) | - | -0.037 | 0.970 |
| TIANJIN | Number  | 51 | 61 | 106 |  |  |
| Age | 34.94±8.27 | 34.67±8.84 | 34.92±10.98 | 0.015 | 0.986 |
|  | Male/Female | 31/20 | 29/32 | 48/58 | 3.445 | 0.181 |
|  | TIV | 1531.61±157.82 | 1506.94±178.58 | 1489.63±150.03 | 1.193 | 0.305 |
|  | Total CPZ | 400 (150. 575) | 400 (300, 587.5) | - | -0.802 | 0.422 |
|  | Age when SZ was firstly diagnosed | 21.5 (19, 28.25) | 22.5 (18, 28) | - | -0.250 | 0.802 |
|  | Age when symptoms firstly appeared | 21 (19, 27) | 22 (18, 26) | - | -0.542 | 0.588 |
|  | Age when therapy was firstly received | 20.5 (18.25, 27.5) | 23 (18.5, 28.5) | - | -0.286 | 0.775 |
|  | Ill duration | 9 (4, 14) | 10 (5.5, 19.5) | - | -1.187 | 0.235 |
|  | Education level  | 4 (3, 7) | 4 (4, 9) | - | -0.149 | 0.881 |
| GUANGZHOU | Number  | 41 | 10 | 29 |  |  |
| Age | 26.61±7.08 | 22.50±4.35 | 25.32±5.87 | 1.720 | 0.186 |
|  | Male/Female | 22/19 | 8/2 | 159/98 | 2.200 | 0.319 |
|  | TIV | 1416.61±195.42 | 1341.80±287.95 | 1452.29±130.30 | 1.286 | 0.282 |

**Supplementary Table 2. The demographic and clinical characteristics of the subjects clustered using K-means method based on PANSS items scores.**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Sites  |  | Subtype 1 | Subtype 2 | HCs | F/Z/χ2 value | P value |
| BrainGluschi | Number  | 23 | 48 | 44 |  | - |
| Age | 33.65±12.93 | 34.54±12.49 | 34.32±10.58 | 0.044 | 0.957 |
|  | Male/Female | 21/2 | 42/6 | 38/6 | 0.325 | 0.934 |
|  | TIV | 1554.86±169.52 | 1497.76±128.81 | 1527.29±117.85 | 1.505 | 0.226 |
|  | Total CPZ | 300 (240, 562) | 262.5 (103.75, 502.5) | - | -0.916 | 0.360 |
|  | Age when SZ was firstly diagnosed | 19 (17.75, 22)  | 20 (17, 24) | - | -0.603 | 0.547 |
|  | Age when symptoms firstly appeared | 18 (16, 21.75) | 19 (16.25, 24) | - | -0.609 | 0.542 |
|  | Age when therapy was firstly received | 20 (17, 25) | 20 (16.5, 26) | - | -0.199 | 0.842 |
|  | Ill duration | 9 (2, 20) | 8 (3, 18) | - | -0.285 | 0.776 |
|  | Education level  | 3 (2, 4) | 3 (3, 4) | - | -0.374 | 0.709 |
| COBRE | Number  | 29 | 51 | 78 |  | - |
| Age | 38.14±12.77 | 36.55±12.44 | 37.12±10.85 | 0.169 | 0.844 |
|  | Male/Female | 23/6 | 41/10 | 56/22 |  | 0.480 |
|  | TIV | 1467.76±137.99 | 1511.68±141.61 | 1497.81±159.53 | 0.794 | 0.454 |
|  | Total CPZ | 400 (200, 618.75) | 300 (132, 600) | - | -1.078 | 0.281 |
|  | Age when SZ was firstly diagnosed | 19 (17, 25) | 19.5 (17, 24) | - | -0.339 | 0.734 |
|  | Age when symptoms firstly appeared | 18 (16, 23) | 18 (16, 24) | - | -0.653 | 0.514 |
|  | Age when therapy was firstly received | 20 (17, 25) | 21 (18, 28) | - | -0.484 | 0.628 |
|  | Ill duration | 11 (4.5, 32.5) | 12 (6, 19) | - | -0.160 | 0.873 |
|  | Education level  | 3 (3, 4) | 4 (3, 4) | - | -0.794 | 0.427 |
| TIANJIN | Number  | 49 | 63 | 106 |  |  |
| Age | 34.55±8.94 | 34.98±8.30 | 34.92±10.98 | 0.031 | 0.970 |
|  | Male/Female | 26/23 | 34/29 | 48/58 |  | 0.471 |
|  | TIV | 1500.24±154.09 | 1532.13±179.96 | 1489.63±150.03 | 1.412 | 0.246 |
|  | Total CPZ | 400 (300, 590) | 400 (225, 572.5) | - | -0.053 | 0.958 |
|  | Age when SZ was firstly diagnosed | 23 (17.75, 28.25) | 22 (18.75, 27.25) | - | -0.063 | 0.950 |
|  | Age when symptoms firstly appeared | 21 (18, 28) | 22 (19, 26) | - | -0.181 | 0.856 |
|  | Age when therapy was firstly received | 24 (18, 30) | 22 (18.75, 26.25) | - | -0.89 | 0.407 |
|  | Ill duration | 9 (4, 19) | 10 (6, 15) | - | -0.447 | 0.690 |
|  | Education level  | 5 (4, 9) | 4 (3, 7) | - | -0.704 | 0.481 |
| GUANGZHOU | Number  | 21 | 30 | 29 |  |  |
| Age | 27.57±8.15 | 24.57±5.48 | 25.31±5.87 | 1.412 | 0.250 |
|  | Male/Female | 10/11 | 20/10 | 17/12 | 1.850 | 0.397 |
|  | TIV | 1402.78±186.40 | 1401.35±236.64 | 1452.29±130.29 | 0.646 | 0.527 |

**Supplementary Table 3. The demographic and clinical characteristics of the subjects clustered using Hierarchical method based on GMV data.**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Sites  |  | Subtype 1 | Subtype 2 | HCs | F/Z/χ2 value | P value |
| BrainGluschi | Number  | 29 | 42 | 44 |  | - |
| Age | 33.66±11.69 | 34.67±13.23 | 34.32±10.58 | 0.063 | 0.939 |
|  | Male/Female | 25/4 | 38/4 | 38/6 | 0.537 | 0.818 |
|  | TIV | 1528.03±139.85 | 1508.12±148.84 | 1527.29±117.85 | 0.276 | 0.759 |
|  | Total CPZ | 225 (56.8, 425) | 300 (200, 600) | - | -1.885 | 0.059 |
|  | Age when SZ was firstly diagnosed | 18.5 (17, 21.75)  | 20 (18, 25.5) | - | -1.624 | 0.104 |
|  | Age when symptoms firstly appeared | 18 (16, 20.5) | 20 (16.75, 25) | - | -1.960 | 0.050 |
|  | Age when therapy was firstly received | 19 (16, 24) | 21 (17.75, 29) | - | -1.870 | 0.062 |
|  | Ill duration | 9.5 (2.5, 22.75) | 7.5 (2.75, 16.25) | - | -0.756 | 0.449 |
|  | Education level  | 3 (3, 4) | 3 (2, 4) | - | -0.251 | 0.801 |
| COBRE | Number  | 30 | 50 | 78 |  | - |
| Age | 35.43±12.02 | 38.14±12.79 | 37.12±10.85 | 0.500 | 0.607 |
|  | Male/Female | 26/4 | 38/12 | 56/22 | 2.624 | 0.274 |
|  | TIV | 1520.09±150.12 | 1481.17±134.74 | 1497.81±159.53 | 0.632 | 0.533 |
|  | **Total CPZ** | **225 (80, 525)** | **450 (220, 618.75)** | **-** | **-2.224** | **0.026\*** |
|  | Age when SZ was firstly diagnosed | 18.5 (17, 24) | 20 (17, 25) | - | -0.546 | 0.585 |
|  | Age when symptoms firstly appeared | 18 (17, 23.25) | 18.5 (16, 24) | - | -0.041 | 0.967 |
|  | **Age when therapy was firstly received** | **19 (17, 25)** | **22 (19, 29)** | **-** | **-2.077** | **0.038\*** |
|  | Ill duration | 15 (5.75, 19) | 11 (6, 22.75) | - | -0.249 | 0.804 |
|  | Education level  | 3 (3, 4) | 4 (3, 4) | - | -0.074 | 0.941 |
| TIANJIN | Number  | 26 | 86 | 106 |  |  |
| Age | 35.04±8.99 | 34.72±8.46 | 34.92±10.98 | 0.015 | 0.986 |
|  | Male/Female | 14/12 | 46/40 | 48/58 | 1.498 | 0.494 |
|  | TIV | 1556.31±171.38 | 1506.65±167.76 | 1489.63±150.03 | 1.833 | 0.162 |
|  | Total CPZ | 400 (125, 570) | 400 (265, 587.5) | - | -0.675 | 0.499 |
|  | Age when SZ was firstly diagnosed | 20 (19, 28) | 23 (18, 28) | - | -0.458 | 0.647 |
|  | Age when symptoms firstly appeared | 20 (19, 23) | 22 (18, 27) | - | -0.438 | 0.661 |
|  | Age when therapy was firstly received | 20 (19, 29.5) | 23 (18, 28) | - | -0.373 | 0.709 |
|  | Ill duration | 10 (5.5, 16.5) | 9 (5, 15) | - | -0.270 | 0.787 |
|  | Education level  | 5 (3, 6.75) | 4 (4, 9) | - | -0.032 | 0.975 |
| GUANGZHOU | Number  | 31 | 20 | 29 |  |  |
| Age | 27.39±7.39 | 23.35±4.99 | 25.31±5.87 | 2.535 | 0.086 |
|  | Male/Female | 19/12 | 11/9 | 17/12 | 0.199 | 0.958 |
|  | TIV | 1428.66±209.69 | 1360.54±222.82 | 1452.29±130.29 | 1.452 | 0.240 |

**Supplementary Table 4. The demographic and clinical characteristics of the subjects clustered using Hierarchical method based on PANSS items scores.**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Sites  |  | Subtype 1 | Subtype 2 | HCs | F/Z/χ2 value | P value |
| BrainGluschi | Number  | 28 | 43 | 44 |  | - |
| Age | 33.61±12.55 | 34.67±12.68 | 34.32±10.58 | 0.069 | 0.934 |
|  | Male/Female | 26/2 | 37/6 | 38/6 | 0.834 | 0.719 |
|  | TIV | 1559.50±157.78 | 1488.09±129.48 | 1527.29±117.85 | 2.547 | 0.083 |
|  | Total CPZ | 300 (240, 600) | 200 (100, 467.5) | - | -1.255 | 0.209 |
|  | Age when SZ was firstly diagnosed | 19 (17, 22)  | 20 (17, 24) | - | -0.763 | 0.446 |
|  | Age when symptoms firstly appeared | 18 (16, 21) | 19 (16, 24) | - | -0.841 | 0.400 |
|  | Age when therapy was firstly received | 20 (16.5, 25) | 21 (17, 27) | - | -0.400 | 0.689 |
|  | Ill duration | 9 (2, 18) | 8 (3, 18) | - | -0.169 | 0.866 |
|  | Education level  | 3 (3, 4) | 3 (3, 4) | - | -0.130 | 0.897 |
| COBRE | Number  | 35 | 45 | 78 |  | - |
| Age | 37.77±12.66 | 36.62±12.50 | 37.12±10.85 | 0.094 | 0.910 |
|  | Male/Female | 28/7 | 36/9 | 56/22 | 0.456 | 0.510 |
|  | TIV | 1489.52±139.93 | 1500.62±143.29 | 1497.81±159.53 | 0.057 | 0.945 |
|  | Total CPZ | 300 (200, 600) | 300 (150, 600) | - | -0.554 | 0.580 |
|  | Age when SZ was firstly diagnosed | 18.5 (17, 25) | 20 (17, 24) | - | -0.758 | 0.449 |
|  | Age when symptoms firstly appeared | 18 (16, 22.5) | 19 (16, 24) | - | -0.766 | 0.444 |
|  | Age when therapy was firstly received | 20 (17, 25) | 22 (18, 28.25) | - | -0.842 | 0.400 |
|  | Ill duration | 11 (5, 32) | 12 (6.5, 19) | - | -0.053 | 0.957 |
|  | Education level  | 4 (3, 4) | 4 (3, 4) | - | -0.057 | 0.955 |
| TIANJIN | Number  | 57 | 55 | 106 |  |  |
| Age | 35.33±9.33 | 34.24±7.70 | 34.92±10.98 | 0.179 | 0.836 |
|  | Male/Female | 32/25 | 28/27 | 48/58 | 1.803 | 0.404 |
|  | TIV | 1510.86±161.43 | 1525.76±177.96 | 1489.63±150.03 | 0.983 | 0.376 |
|  | Total CPZ | 400 (300, 500) | 400 (225, 600) | - | -0.148 | 0.882 |
|  | Age when SZ was firstly diagnosed | 22 (18, 28) | 24 (18.5, 28) | - | -0.666 | 0.505 |
|  | Age when symptoms firstly appeared | 21 (19, 26.75) | 21.5 (18, 26.25) | - | -0.055 | 0.956 |
|  | Age when therapy was firstly received | 22 (18, 29) | 25 (18.75, 28) | - | -0.323 | 0.747 |
|  | Ill duration | 12 (4, 19.5) | 9 (5, 14) | - | -0.911 | 0.362 |
|  | Education level  | 5 (4, 8.5) | 4 (3, 9) | - | -0.348 | 0.728 |
| GUANGZHOU | Number  | 26 | 25 | 29 |  |  |
| Age | 26.08±7.12 | 25.52±6.58 | 25.31±5.87 | 0.099 | 0.905 |
|  | Male/Female | 11/15 | 19/6 | 17/12 | 5.970 | 0.051 |
|  | TIV | 1381.62±175.52 | 1423.07±252.22 | 1452.29±130.29 | 0.955 | 0.389 |