**Supplementary Figure 4:** Forest plots of the variance explained (on the liability scale) of European ancestry samples divided according to postulated sources of heterogeneity.

Samples were sub-grouped according to the following criteria:

1) Case definition: schizophrenia (SCZ), schizophrenia or schizoaffective disorder (SCZ/Schizoaffective), schizophrenia spectrum disorder

2) Screened or unscreened controls for schizophrenia or other psychoses

3) Recruitment setting: Cases recruited from Mixed (ie Community or Hospital setting), Community only, or Hospital plus ascertainment for clozapine treatment.

4) Diagnostic strategy: Consensus diagnosis, research diagnostic interview, review of medical records, mixed strategy **(see Supplementary Note**)

For each sample, the mid-point of the box represents the point estimate of the variance explained (R2) and the horizontal line the 95% confidence intervals for that estimate. The shaded area of the box is proportional to the corresponding study weight, i.e. the cohort's contribution to the overall estimate of variance explained (R2) estimated with the DerSimonian & Laird random effects model. Heterogeneity is denoted by the I2 statistic; measures above 75% indicating considerable heterogeneity among the samples within each subgroup. Unshaded diamonds represent the mean estimate for studies in the relevant section of the plot. The overall mean of the European sample is represented by the dashed line.

Details of all samples are given in the **Sample Supplementary note**.