Figure S2: Pooled SMD about psychiatric symptoms which assessed with different scales Experimental Control Std. Mean Difference Std. Mean Difference Study or Subgroup Mean SD Total Mean SD Total Weight IV, Random, 95% CI IV, Random, 95% CI **8.1.1 PANSS** Huang YB, et al 2019 23.96 2.87 80 32.81 3.95 80 14.7% -2.55 [-2.97, -2.13] Liu L. et al 2018 42.22 6.52 45 51.28 6 49 14.4% -1.38 [-1.84, -0.92] Liu SS. et al 2017 43.62 5.93 -1.23 [-1.67, -0.79] 58 51.26 6.53 14.6% Xiang JF. et al 2019 42.81 5.88 45 50.35 6.21 14.5% -1.24 [-1.69, -0.78] Yang J. et al 2017 42.68 9.02 41 53.39 10.55 14.4% -1.08 [-1.55, -0.62] Ying C. et al 2020 46.07 15.07 28 59.46 20.57 13.4% -0.74 [-1.31, -0.17] 23 Subtotal (95% CI) 297 274 86.0% -1.38 [-1.89, -0.87] Heterogeneity:  $Tau^2 = 0.35$ ;  $Chi^2 = 36.68$ , df = 5 (P < 0.00001);  $I^2 = 86\%$ Test for overall effect: Z = 5.27 (P < 0.00001) 8.1.2 BPRS Shen YR, et al 2016 21.18 3.15 81 28.42 4.02 14.0% -2.11 [-2.62, -1.61] Subtotal (95% CI) 81 14.0% -2.11 [-2.62, -1.61] Heterogeneity: Not applicable Test for overall effect: Z = 8.19 (P < 0.00001) Total (95% CI) 378 304 100.0% -1.48 [-1.96, -1.01] Heterogeneity:  $Tau^2 = 0.35$ ;  $Chi^2 = 42.57$ , df = 6 (P < 0.00001);  $I^2 = 86\%$ Test for overall effect: Z = 6.15 (P < 0.00001) Favours [experimental] Favours [control] Test for subgroup differences:  $Chi^2 = 3.96$ , df = 1 (P = 0.05),  $I^2 = 74.8\%$