**SUPPLEMENTARY MATERIAL**

**GROUP investigators**

Therese van Amelsvoort1, Agna Bartels-Velthuis2, Richard Bruggeman2,3, Lieuwe de Haan4,5, Frederike Schirmbeck4,5, Claudia Simons1,6, Jim van Os7,8

1Department of Psychiatry and Neuropsychology, Maastricht, the Netherlands; 2University Centre for Psychiatry, University Medical Centre Groningen, Groningen, the Netherlands; 3Department of Clinical and Developmental Neuropsychology, University of Groningen, Groningen, the Netherlands; 4Department of Psychiatry, Academic Medical Centre, Amsterdam, the Netherlands; 5Arkin Institute for Mental Health, Amsterdam, the Netherlands; 6GGzE Institute for Mental Health Care, Eindhoven, the Netherlands; 7Department of Psychiatry, University Medical Centre Utrecht, Utrecht, the Netherlands; and 8Department of Psychosis Studies, King’s College London, London, UK

**NEDEN investigators**

Tim Amos1, Max Birchwood2, Linda Everard3, Nick Freemantle4

1Avon and Wiltshire Mental Health Partnership National Health Service Trust, Bristol, UK; 2Unit of Mental Health and Wellbeing, University of Warwick, UK; 3Birmingham and Solihull Mental Health National Health Service Foundation Trust, Birmingham, UK; and 4Department of Primary Care and Population Health, University College London, London, UK

**Table S1. Comparisons on model variables between complete and incomplete cases in the GROUP sample (*N* = 1013).**

 Model variable CC (*n* = 871) IC (*n* = 142) Test statistic (*df*) *p*

 Age, mean (SD) in years 27.3 (7.1) 25.2 (6.3) *t* (1011) = 3.31 .001

 Male 673 (77) 117 (82) *χ2* (1) = 1.87 ns

 Completed secondary schoola 753 (86) 106 (80) *χ2* (1) = 3.53 ns

 Use of illicit substances 602 (69) 89 (75) *χ2* (1) = 1.60 ns

 Violent behaviour 179 (21) 4 (20) - -

CC, complete cases; IC, incomplete cases; *df*, degrees of freedom; SD, standard deviation; ns, nonsignificant.

Data are *n* (%), unless otherwise stated. Due to missing data, the number of incomplete cases varies by model variable.

**Table S2. Comparisons on model variables between complete and incomplete cases in the NEDEN sample (*N* = 1027).**

 Model variable CC (*n* = 921) IC (*n* = 106) Test statistic (*df*) *p*

 Age, mean (SD) in years 22.8 (4.8) 22.1 (5.3) *t* (1008) = 1.35 ns

 Male 639 (69) 70 (66) *χ2* (1) = 0.50 ns

 Completed secondary schoola 811 (88) 61 (80) *χ2* (1) = 3.89 .049

 Use of illicit substances 589 (64) 46 (67) *χ2* (1) = 0.21 ns

 Violent behaviour 204 (22) 15 (23) *χ2* (1) = 0.03 ns

CC, complete cases; IC, incomplete cases; *df*, degrees of freedom; SD, standard deviation; ns, nonsignificant.

Data are *n* (%), unless otherwise stated. Due to missing data, the number of incomplete cases varies by model variable.

**Table S3. Prevalence and risk of violent behaviour by severity of use of alcohola and different categories of illicit substances in the GROUP sample (*N* = 871).**

 *n* (%) aOR (95% CI)b

 Substance category PU NPU NU PU *vs* NPU or NU PU*vs* NU NPU *vs* NU

 Alcohol 83 (24) 74 (19) 21 (16) 1.5 (1.1-2.1) 1.8 (1.0-3.1) 1.3 (0.7-2.2)

 Cannabis 90 (24) 45 (21) 44 (15) 1.3 (0.9-1.9) 1.5 (1.0-2.4) 1.4 (0.8-2.2)

 Stimulants 43 (30) 31 (23) 98 (17) 1.8 (1.2-2.7) 2.0 (1.3-3.0) 1.4 (0.9-2.3)

 Depressants 15 (33) 18 (24) 134 (19) 1.9 (1.0-3.7) 2.0 (1.0-3.9) 1.3 (0.7-2.4)

 Hallucinogens 12 (27) 41 (30) 113 (17) 1.3 (0.7-2.7) 1.6 (0.8-3.2) 2.0 (1.3-3.1)

 Other 28 (29) 42 (28) 100 (17) 1.5 (0.9-2.4) 1.9 (1.1-3.1) 1.9 (1.2-2.9)

aOR, adjusted odds ratio; CI, confidence interval; PU, problematic use; NPU, nonproblematic use; NU, no use.

Due to missing data, the total number of patients varies by substance category.

aAlcohol was used as a positive control.

bAdjusted for age, sex and educational level.