Data supplement for Kendler et al, Prediction of Onset of Substance-Induced Psychotic Disorder and Its Progression to Schizophrenia in a Swedish National Sample. Am J Psychiatry (doi: 10.1176/appi.ajp.2019.18101217)

## Definition of Drug Abuse

Based on information in the medical, criminal and prescribed drug Register we defined Drug Abuse (DA). In the Swedish medical registries DA was identified from primary and secondary ICD codes (ICD8: Drug dependence (304); ICD9: Drug psychoses (292) and Drug dependence (304), Nondependent abuse of drugs (305.1-305.9); ICD10: Mental and behavioral disorders due to psychoactive substance use (F10-F19), except those due to alcohol (F10) or tobacco (F17)); in the Suspicion Register by codes 3070, 5010, 5011, and 5012, that reflect crimes related to DA; and in the Crime Register by references to laws covering narcotics (law 1968:64, paragraph 1, point 6) and drug-related driving offences (law 1951:649, paragraph 4, subsection 2 and paragraph 4A, subsection 2). DA was identified in individuals (excluding those suffering from cancer) in the Prescribed Drug Register who had retrieved (in average) more than four defined daily doses a day for 12 months from either of Hypnotics and Sedatives (Anatomical Therapeutic Chemical (ATC) Classification System N05C and N05BA) or Opioids (ATC: N02A). An individual was considered as registered for DA if he/she fulfilled any of the criteria described above.

## Definition of Alcohol Use Disorder

Alcohol Use Disorders (AUD) was defined in three ways: 1) by ICD codes for main and secondary diagnoses from Swedish medical and Cause of Death registries (the Swedish Hospital Discharge Register, containing all hospitalizations for all Swedish inhabitants from 1973-2012 and the Outpatient Care Register, containing information from all outpatient clinics from 2001 to 2012, the Swedish Cause of Death Register, containing information on all deaths in Sweden from 1963 to 2012) for the following diagnoses: ICD8 and 9: alcohol-related psychiatric disorders (291), alcohol dependence (303), alcohol abuse (305A), alcohol-related polyneuropathy (357F), alcohol-related cardiomyopathy (425F), alcohol-related gastritis (535D), alcoholic fatty liver, alcohol hepatitis, alcoholic cirrhosis, unspecified liver damage caused by alcohol (571A-D), toxic effects of alcohol (980), alcoholism (V79B); ICD10: alcohol6 related psychiatric and behavioral disorders (F10, excluding acute alcohol intoxication: F10.0), rehabilitation of a person with alcohol abuse (Z50.2), guidance and medical advice to a person with alcohol abuse (Z71.4), alcohol-related pseudo-Cushing syndrome (E24.4), alcohol-related degeneration of the nervous system and brain (G31.2), alcohol-related polyneuropathy (G62.1), alcohol-related myopathy (G72.1), alcohol-related cardiomyopathy (I42.6), alcohol-related gastritis (K29.2), liver diseases caused by alcohol (K70.0-K70.9), acute pancreatitis caused by alcohol (K85.2), chronic pancreatitis caused by alcohol (K86.0), treatment of pregnant alcoholic woman (O35.4), toxic effects of alcohol (T51.0-T51.9); 2) by Anatomical

Therapeutic Chemical (ATC) codes in the Prescribed Drug Register (containing all prescriptions in Sweden picked up by patients from July 2005 to 2012): disulfiram (N07BB01), acamprosate (N07BB03), or naltrexone (N07BB04; 3) by registrations of individuals in the Swedish Crime Registers (the Swedish Crime Register included national complete data on all convictions from 1973-2010 and the Swedish suspicion register included national complete data on all individuals strongly suspected of crime from 1998-2010) with at least two convictions of drunk driving (suspicion code 3005, law 1951:649 (paragraph 4 and 4A)) or drunk in charge of a maritime vessel (suspicion code 3201, law 1994:1009 (chapter 20, paragraph 4 and 5)). We insured that we did not count arrests in the suspicion register that described the same event contained in the conviction register.

## Definition of Early Retirement

Early retirement was categorized into a binary variable based on whether or not the individual has received any early retirement prior to a particular year. Early retirement is a combination of several different variables from the Swedish registers. Until year 2002, Early Retirement Pension was paid to people aged 16- 64 and granted when their working capability was deemed to be permanently reduced by at least one quarter due to medical reasons. The early retirement variable is also composed of temporary disability pensions paid to individuals whose working capability was not expected to be permanent but was expected to persist for a considerable time. From 2003, the rules for Early Retirement Pension and Temporary disability pension changed and these types of compensations were changed into sickness compensation (for individuals 30-64) and activity compensation (for individuals 19-29). The qualification rules were similar, but activity compensation was supposed to be limited in time. The determination of early retirement is made by the Swedish Social Insurance Agency.

## Calculation of Familial Risk Scores

We calculated the morbid risk for NAP for the entire Swedish population born 1940-1995. We divided the population into three groups based on the distribution of age at first ANAP registration: individuals in the first quartile (below 25 years) were weighted 0; individuals in the second and third quartile (25-42 years) were weighted 0.5; individuals in the last quartile (above 42 years) were weighted 1. Thereafter, we performed a logistic regression analysis based on information from the Swedish population born between 1940 and 1995, and their relatives (MZ twins, full siblings, half-siblings, mother, father, grandparents, aunts, uncles and cousins). The model outcome variable was NAP and the predictor variables the morbid risk of NAP among the different relative types. The resulting beta weights for the genetic risk score followed quantitative genetic expectations. This model produced a predicted probability for NAP which we called the familial risk score (FRS). This procedure was replicated for DA and AUD. We utilized the FRS for NAP rather than for schizophrenia for two reasons. First, it better captured the liability to SIPD by reflecting the broader vulnerability to psychosis rather than the narrower vulnerability to schizophrenia. Second, the distribution of the FRS for NAP was considerably less skewed than that seen for the RFS for schizophrenia and considerably more informative. The two FRS scores were substantially but not incompletely inter-correlated (+0.58).

**FIGURE S1. A Receiver Operator Characteristic Curve for the Prediction of Substance Induced Psychotic Disorder**

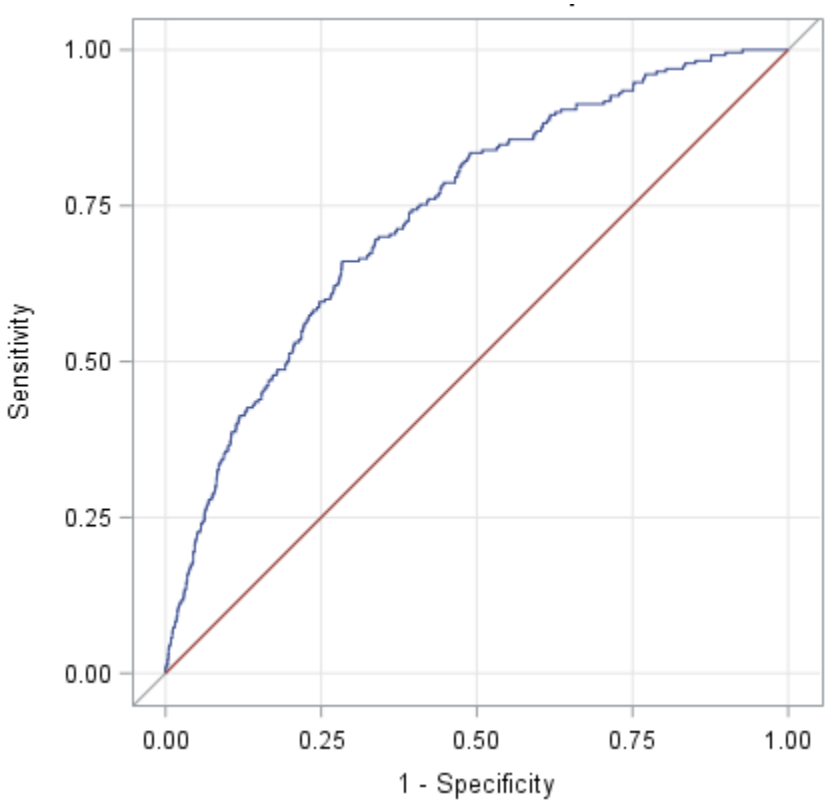

**FIGURE S2. Survival Curves for Schizophrenia in Cases of Substance Induced Psychotic Disorder as a Function of Time Since First Onset Requiring 2 Diagnoses to be Considered a Case of Schizophrenia**

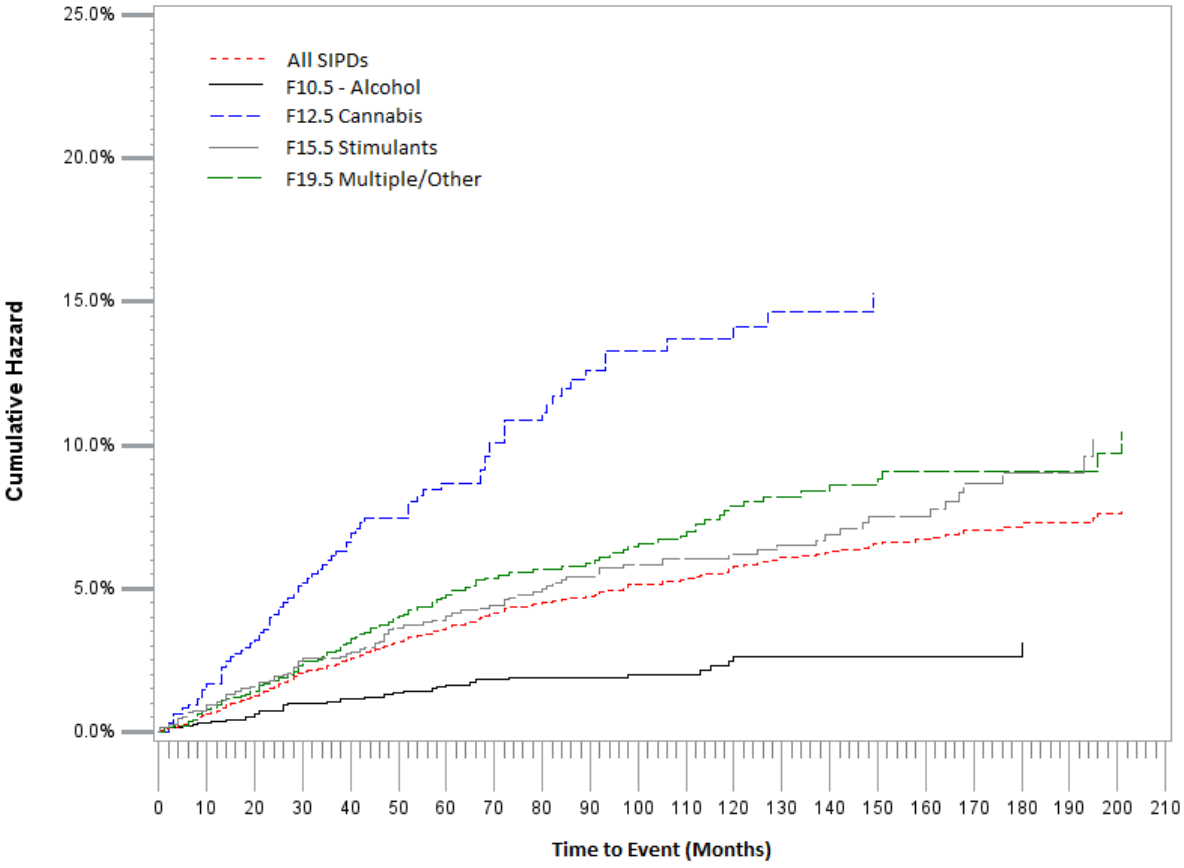

**FIGURE S3. Rates of Receipt of Early Retirement for Cases of Cannabis-Induced Psychotic Disorders Who Did versus Did Not Receive a Subsequent Schizophrenia (SZ) Diagnosis. Dotted Lines Indicate 95% Confidence Intervals**

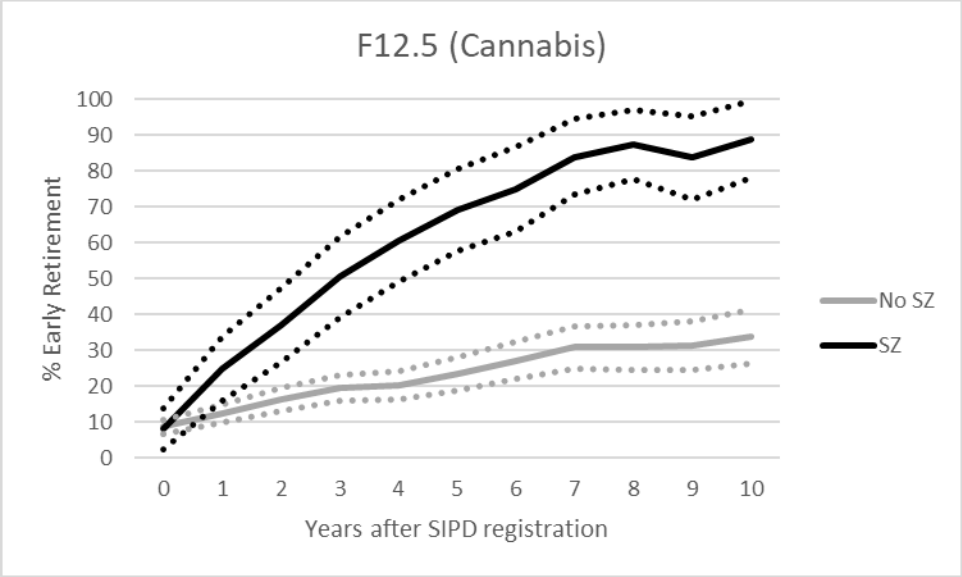

**TABLE S1. Mean Standardized Familial Risk Scores for Drug Abuse (DA), Alcohol Use Disorder (AUD) and Non-Affective Psychosis (NAP) in Cases of Substance-Induced Psychotic Disorder (SIPD) that Did versus Did Not Convert to Schizophrenia Defined as Having at Least Two Diagnoses**

| | Conversion to Schizophrenia | All | F10.5 Alcohol | F12.5 Cannabis | F15.5 Stimulants | F19.5 (Multiple/ Other) | P value across 4 specific forms of SIPD |
|---|---|---|---|---|---|---|---|
| | | | | | | | |
| DA Familial Risk Score | Yes | 1.21<br>0.89; 1.54 | 0.09<br>-0.21; 0.38 | 0.98<br>0.46; 1.49 | 1.73<br>0.98; 2.49 | 1.29<br>0.75; 1.82 | 0.0020 |
| DA Familial Risk Score | No | 1.08<br>1.02; 1.15 | 0.42<br>0.34; 0.50 | 1.28<br>1.08; 1.49 | 1.16<br>1.02; 1.30 | 1.60<br>1.45; 1.74 | <.0001 |
| P-value difference | | 0.1791 | 0.1543 | 0.1010 | 0.0570 | 0.2543 | |
| | | | | | | | |
| AUD Familial Risk Score | Yes | 1.06<br>0.79; 1.34 | 1.16<br>0.40; 1.92 | 0.58<br>0.15; 1.01 | 1.37<br>0.87; 1.87 | 0.97<br>0.53; 1.42 | 0.0029 |
| AUD Familial Risk Score | No | 0.98<br>0.93; 1.03 | 0.90<br>0.82; 0.99 | 0.62<br>0.50; 0.73 | 1.18<br>1.07; 1.30 | 1.09<br>0.99; 1.19 | <.0001 |
| P-value difference | | 0.4834 | 0.0784 | 0.2196 | 0.1019 | 0.1056 | |
| | | | | | | | |
| NAP Familial Risk Score | Yes | 0.55<br>0.32; 0.78 | 0.82<br>0.24; 1.40 | 0.82<br>0.16; 1.48 | 0.93<br>0.06; 1.80 | 0.58<br>0.04; 1.13 | 0.4744 |
| NAP Familial Risk Score | No | 0.35<br>0.28; 0.40 | 0.24<br>0.13; 0.36 | 0.43<br>0.28; 0.58 | 0.37<br>0.27; 0.48 | 0.38<br>0.29; 0.47 | <.0001 |
| P-value difference | | <0.0001 | <0.0001 | 0.0803 | 0.0009 | 0.0315 | |
| The p-values are calculated using a non-parametric approach; the Van der Waerden Scores. | | | | | | | |

**TABLE S2. Mean Standardized Familial Risk Scores for Non-Affective Psychosis (NAP) in Cases of Schizophrenia (defined by two or more diagnoses) With and Without a Prior Diagnosis of Substance-Induced Psychotic Disorder (SIPD)**

| | Genetic Risk Score (ANAP) | Genetic Risk Score (DA) | Genetic Risk Score (AUD) |
|---|---|---|---|
| SZ – No SIDP | 0.92 | 0.04 | 0.21 |
| SZ – SIPD | 0.67 | 0.77 | 0.84 |
| P-value difference | 0.2338 | <0.0001 | <0.0001 |
| | | | |

The p-values (One-sided test) are calculated using a non-parametric approach; the Van der Waerden Scores.