APPENDIX

Table 1 - Definition of phenotypes

|  |  |  |
| --- | --- | --- |
|  | Registers Used | Definition  |
| Alcohol Use Disorder (AUD) | The Swedish Hospital Discharge Register (coverage 1973-2017); Outpatient Care Register (national coverage 2001-2017); Primary Care Registry (Partly coverage from 1999-2017); the Swedish Drug Register (2005-2017); the Swedish Mortality Register, and the Swedish Criminal Register (1973-2017) and the Swedish Suspicion Register (1998-2017)  | Alcohol Use Disorder (AUD) was identified in the Swedish medical and mortality registries by ICD codes: ICD8/9: V79B, 305A, 357F, 571A-D, 425F, 535D, 291, 303, 980; ICD 10: E244, G312, G621, G721, I426, K292, K70, K852, K860, O354, T51, F10); in the Crime Register by codes 3005, 3201, which reflect crimes related to alcohol abuse; in the Suspicion Register by codes 0004, 0005 (Only those individuals with at least two alcohol-related crimes or suspicion of crimes from both Crime Register and Suspicion Register were included); in the Prescribed Drug Register by the drugs disulfiram (Anatomical Therapeutic Chemical (ATC) Classification System N07BB01), acamprosate (N07BB03), and naltrexone (N07BB04). |
| Drug Use Disorder (DUD) | The Swedish Hospital Discharge Register (coverage 1973-2017); Outpatient Care Register (national coverage 2001-2017); Primary Care Registry (Partly coverage from 1999-2017); the Swedish Drug Register (2005-2017); the Swedish Mortality Register, and the Swedish Criminal Register (1973-2017) and the Swedish Suspicion Register (1998-2017) | Drug abuse (DA) was identified in the Swedish medical and mortality registries by ICD codes (ICD8: Drug dependence (304); ICD9: Drug psychoses (292) and Drug dependence (304); 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). |
| Educational attainment (EA) | The longitudinal integration database for health insurance and labor market studies (LISA) from 1990-2017 and The Swedish Census from 1970. | Highest achieved education measured in 1-7 levels that in turn were translated into number of years of education and then standardized with mean 0 and SD 1 by gender and year of birth. In the registers the variable are as follows:1-Pre-high school (7 years)2-High School (9 years)3-Upper Secondary School (11 years)4-Upper Secondary School (12 years)5-Post-secondary education (14 years)6-Post-secondary education (17 years)7- PhD education (21 years) |

**Table 2 Steps for the Calculation of Familial genetic risk score (FGRS):**

The dataset for the calculations includes:

Column1 = Identification number of the proband (Born 1932-1995)

Column2 = Identification number of the relative (1st to 5th degree relatives)

Column3 = Proportion of shared genetic (0.03125 to 1) with the proband

Column4 = Year of Birth of relative

Column5 = Sex of relative

Column6 = Age at registration for trait

Column7 = Age at end of follow-up (2017-12-31 or age at death, or age at emigration whichever came first)

**Step 1:** Using all unique relatives with a registration for the trait we non-parametrically estimated the distribution of *Age at first registration*. The empirical distribution is used to obtain weights for relatives without a registration for the trait, in order to account for the proportion of time-at-risk period they had completed at the end of follow-up. For example, for relatives at age x at end of follow-up the weight corresponds to the proportion of relatives registered for the trait that had been registration at age x. For relatives born prior to 1958 we subtracted age at end of follow-up with the following formula: 1958 - Year of birth of relative. This modification was done in order to control for registration effects (i.e, most registers start in 1973 suggesting that relatives from early birth cohorts do not have the possibility to be registered at younger ages). Note that all relatives with the trait is weighted one.

**Step 2:** Transform the binary variable (trait yes/no) into a z-score based on the threshold for each trait. The underlying liability of the individual is not assessable. Instead we estimated the mean of the underlying liability to obtain sex and birth decade specific Z-scores for relatives with the trait registration and relatives without the trait. We generate n random numbers from a N(0, 1) distribution and estimate the mean for relatives registered with the trait (i.e., mean of the observations above the threshold) and for relatives without a registrations (i.e., mean of all observation below the threshold). The thresholds are calculated for each decade of birth and sex.

**Step 3**: Correct for cohabitation effects. To estimate the cohabitation effect, we created a database with all individuals in the Swedish population born in Sweden 1955-1990. We also included the number of years, during ages 0-15, that individuals resided in the same household as their biological father. We thereby were able to define two kinds of families i) “not-lived-with” father families (offspring never resided a maximum of 1 year in the same household as their biological father); ii) “lived-with” father (offspring resided a minimum of 13 year in the same household as their biological father. We performed a logistic regression model with the binary trait in offspring as outcome and the binary trait in father, type of father, and their interaction as predictors. We used the interaction term as the difference of effect between genes only and genes + environment. The same approach was performed for half-siblings. The following interaction terms are used in the calculations:

|  |  |  |
| --- | --- | --- |
|  | Parent/Children  | Siblings |
| AUD | 0.99 | 0.69 |
| DUD | 0.92 | 0.52 |
| Education | 0.82 | 0.73 |

**Step 4:** Calculate the product for each relative using the four components:

1. Z-score (reflecting sex and year of birth adjusted rates)
2. Weight (reflecting the proportion of risk period they had completed)
3. Cohabitation effects
4. Proportion of shared genetic (0.03125 to 1) with the proband

**Step 5:** Average the product calculated in step 4 across all relatives to a proband

**Step 6**: Correct for the number of relatives. We multiplied the results from step 5 with a shrinkage factor. Shrinkage factor (SF): B/(B+A/C). It produces more shrinkage if B and C are small and A is large.

1. the variance of the z-score of the trait across all relatives,
2. the variance in the mean z-score across all probands,
3. the weighted number of relatives for each proband (sum of Column 3 across each proband).

**Step 7:** Correct for difference by year of birth and county differences. There are 21 counties in Sweden. For each proband we used the county they had resided in during the maximum number of years (measured from 1969 and onwards) We standardized the risk score by year of birth and county of the proband into a z-score with mean 0 and SD 1. This was then used as the FGRS in the analyses.

The code from the calculation of the FGRS is available from the authors on request.

Figure 1 Mean FGRS/GAEA for individuals with AUD and DUD across Northern and Southern part of Sweden



**Table 3 Correlation with the FGRS used in the manuscript**

|  |  |  |
| --- | --- | --- |
|  | DUD GRS | AUD GRS |
| GRS(a) - 1st degree relatives | 0.763 | 0.812 |
| GRS(b) - no age correction | 0.995 | 0.806 |
| GRS(c) - no cohabitation correction | 0.980 | 0.992 |
| GRS(d) - no weighting for # relatives | 0.971 | 0.952 |
| GRS(e) -std by YoB only | 0.990 | 0.992 |
| GRS(f) - std by geography only | 0.951 | 0.969 |
| GRS(g) - std only by entire sample | 0.946 | 0.962 |
|  |  |  |

Table 4

Correlation with the FGRS reported in ms

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | FGRS based on 1st-4th degree relatives | FGRS based on 1st-3rd degree relatives | FGRS based on 1st-2nd degree relatives | FGRS based on 1st degree relatives |
| AUD | 0.997 | 0.98 | 0.93 | 0.79 |
| DUD | 0.997 | 0.97 | 0.90 | 0.74 |
| EDU | 0.999 | 0.99 | 0.95 | 0.86 |

Table 5

Polychoric Correlation between AUD/DUD and FGRSAUD, FGRSDUD and FGRSEDU

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| FGRS | Outcome | FGRS based on 1st-5th degree relatives | FGRS based on 1st-4th degree relatives | FGRS based on 1st-3rd degree relatives | FGRS based on 1st-2nd degree relatives | FGRS based on 1st degree relatives |
| AUD | AUD | 0.26 (0.00) | 0.26 (0.00) | 0.26 (0.00) | 0.25 (0.00) | 0.19 (0.00) |
| EDU | AUD | 0.18 (0.00) | 0.18 (0.00) | 0.19 (0.00) | 0.19 (0.00) | 0.19 (0.00) |
|  |  |  |  |  |  |  |
| DUD | DUD | 0.26 (0.00) | 0.26 (0.00) | 0.26 (0.00) | 0.24 (0.00) | 0.21 (0.00) |
| EDU | DUD | 0.20 (0.00) | 0.20 (0.00) | 0.20 (0.00) | 0.20 (0.00) | 0.21 (0.00) |

Table 6

Main results using the FGRSs calculated with relatives with different degree of relatedness

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **AUD** | Reported in MS | FGRS based on 1st-4th degree relatives | FGRS based on 1st-3rd degree relatives | FGRS based on 1st-2nd degree relatives | FGRS based on 1st degree relatives |
| FGRS | 6.32 (6.21; 6.43) | 6.39 (6.28; 6.50) | 6.52 (6.41; 6.63) | 6.62 (6.51; 6.73) | 6.42 (6.31; 6.53) |
| GAEA | -2.90 (2.83; 2.97) | -2.94 (-3.01; -2.87) | -3.03 (-2.96; -3.10) | -3.16 (-3.23; -3.09) | -3.41 (-3.49; -3.33) |
| FGRS\* GAEA  | -1.93 (2.03; -1.83) | -2.03 (-2.13; -1.93) | -2.42 (-2.53; -2.31) | -2.75 (-2.87; -2.63) | -2.60 (-2.73; -2.47) |
|  |  |  |  |  |  |
| **DUD** |  |  |  |  |  |
| FGRS | 4.65 (4.56, 4.74) | 4.69 (4.60; 4.78) | 4.77 (4.67; 4.87) | 4.72 (4.62; 4.82) | 4.53 (4.42; 4.64) |
| GAEA | -2.08 (2.13; 2.03) | -2.11 (-2.16; -2.06) | -2.20 (-2.26; -2.14) | -2.30 (-2.36; -2.24) | -2.46 (-2.52; -2.40) |
| FGRS\* GAEA  | -1.58 (1.50; 1.66) | -1.67 (-1.76; -1.76) | -2.08 (-2.18; -1.98) | -2.41 (-2.52; -2.30) | -2.35 (-2.47; -2.23) |
|  |  |  |  |  |  |

Table 7 - Sensitivity analysis:

1. Outcome variable: Time to first AUD registration. Main exposure variables: FGRS AUD and GAEA. Restriction: Individuals with DUD are excluded. 69% of original AUD cases
2. Outcome variable: Time to first DUD registration. Main exposure variables: FGRS DUD and GAEA. Restriction: Individuals with AUD are excluded. 51% of original DUD cases
3. Outcome variable: Time to first AUD/DUD registration. Main exposure variables: FGRS AUD and GAEA. Restriction: Individuals with only AUD and only DUD are excluded.
4. Outcome variable: Time to first AUD/DUD registration. Main exposure variables: FGRS DUD and GAEA. Restriction: Individuals with only AUD and only DUD are excluded

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Model i | Model ii | Model iii | Model iv |
| Familial Genetic Risk Score (FGRS) | 3.86 (3.77; 3.95) | 1.49 (1.42; 1.56) | 1.57 (1.50; 1.64) | 1.49 (1.42; 1.42) |
| Genetic Aptitude for Educational Attainment (GAEA) | -1.79 (-1.85; -1.73) | -0.63 (-0.67; -0.59) | -0.60 (-0.64; -0.56) | -0.63 (-0.67; -0.59) |
| Year of Birth | 0.08 (0.07; 0.09) | 0.35 (0.34; 0.36) | -0.08 (-0.08; -0.07) | -0.08 (-0.08; -0.07) |
| Male vs Female | 7.16 (7.04; 7.28) | 1.10 (1.01; 1.19) | 1.03 (0.95; 1.11) | 1.03 (0.95; 1.11) |
| FGRS\* GAEA  | -0.98 (-1.06; -0.90) | -0.54 (-0.61; -0.47) | -0.53 (-0.59; -0.46) | -0.54 (-0.61; -0.47) |
|  |  |  |  |  |

Figure 2 - Figures illustrating the sensitivity analysis above (from upper left to lower right: Models i through iv)

Table 8 - Results using Cox proportional Hazard models (multiplicative interaction)

|  |  |  |
| --- | --- | --- |
|  | AUD | DUD |
| Familial Genetic Risk Score (FGRS) | 1.40 (1.40; 1.41) | 1.36 (1.35; 1.38) |
| Genetic Aptitude for Educational Attainment (GAEA) | 0.73 (0.73; 0.74) | 0.69 (0.68; 0.69) |
| Year of Birth | 1.01 (1.01; 1.01) | 1.07 (1.07; 1.07) |
| Male vs Female | 2.54 (2.51; 2.58) | 1.79 (1.76; 1.82) |
| FGRS\* GAEA  | 1.02 (1.01; 1.02) | 1.02 (1.01; 1.03) |
|  |  |  |

Figure 3a - Model A using a multiplicative interaction (i.e., Cox proportional Hazards model) – (AUD) (Reference is FGRS AUD: -1 and GAEA = +2)

Figure 3b Model A using a multiplicative interaction (i.e., Cox proportional Hazards model) – (DUD) (Reference is FGRS DUD: -1 and GAEA = +2)