Psychological Medicine (2013), **43**, 2169–2178. © Cambridge University Press 2013
doi:10.1017/S0033291713000081

ORIGINAL ARTICLE

A latent class analysis of drug abuse in a national Swedish sample

K. S. Kendler[1,2,3]*, H. Ohlsson[4], K. Sundquist[4,5] and J. Sundquist[4,5]

[1] *Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA, USA*
[2] *Department of Psychiatry, Virginia Commonwealth University, Richmond VA, USA*
[3] *Department of Human and Molecular Genetics, Virginia Commonwealth University, Richmond, VA, USA*
[4] *Center for Primary Health Care Research, Lund University, Malmö, Sweden*
[5] *Stanford Prevention Research Center, Stanford University School of Medicine, Stanford, CA, USA*

Background. Drug abuse (DA) is a clinically heterogeneous syndrome. Using medical, legal, death and pharmacy records covering the entire population of Sweden, could we uncover meaningful subtypes of DA?

Method. We performed a latent class analysis (LCA) on all individuals in Sweden born 1950–1993 who were registered with DA or its consequences ($n = 192\,501$) and then validated these classes using demographics, patterns of co-morbidity with alcohol use disorder (AUD), non-DA crime and psychiatric illness, and the pattern of aggregation and co-aggregation in sibling pairs.

Results. The best-fit LCA had six classes: (1) low-frequency pure criminal, (2) high-frequency medical criminal, (3) low-frequency pure medical, (4) high-frequency medical, (5) prescription and (6) death. Each class had a distinct pattern of demographic features and co-morbidity and aggregated within sibling pairs with at least moderate specificity. For example, class 2 was characterized by early age at registration, low educational attainment, high male preponderance, high rates of AUDs, strong resemblance within sibling pairs [odds ratio (OR) 12.6] and crime and the highest risk for DA in siblings (20.0%). By contrast, class 5 had a female preponderance, late age at registration, low rates of crime and AUDs, high rates of psychiatric illness, high familiality within sibling pairs (OR 14.7) but the lowest observed risk for DA in siblings (8.9%).

Conclusions. DA as assessed by public records is a heterogeneous syndrome. Familial factors contribute substantially to this heterogeneity. Advances in our understanding of etiological processes leading to DA will be aided by a consideration of this heterogeneity.

Received 2 October 2012; Revised 10 December 2012; Accepted 3 January 2013; First published online 1 February 2013

Key words: Drug abuse, latent class analysis, siblings, Sweden.

Introduction

Drug abuse (DA) is a clinically heterogeneous syndrome with affected individuals frequently differing across several dimensions including the specific abuse and dependence symptoms displayed, the kinds of psychoactive drugs consumed, the route of ingestion, the presence or absence of psychiatric co-morbidity, and the psychosocial correlates and consequences of the abuse (Babor & Caetano, 2006). Given this wide heterogeneity, it is of obvious interest to ask whether more homogeneous subtypes of DA could be defined that might be helpful in future research and clinical care.

* Address for correspondence: K. S. Kendler, M.D., Virginia Institute for Psychiatric and Behavioral Genetics of VCU, Box 980126, Richmond, VA 23298-0126, USA.
(Email: kendler@vcu.edu)

As outlined by Basu *et al*. (2004), many prior attempts have been made to develop typologies of drug abuse. These efforts can be usefully divided into those that classified individuals on single variables (e.g. sex, family history, age at onset or presence/absence of antisocial personality disorder) *versus* those that used multivariate methods (e.g. Wilkinson *et al*. 1987; Alterman & Cacciola, 1991; Byqvist & Olsson, 1998; Babor & Caetano, 2006). Not surprisingly, Basu *et al*. (2004) suggest that multivariate methods are more useful and potentially valid.

One powerful and objective multivariate statistical method to examine this heterogeneity is latent class analysis (LCA). Prior LCAs of DA have been performed in clinical populations (Schwartz *et al*. 2010; Kuramoto *et al*. 2011) and in community samples (Lynskey *et al*. 2006; Agrawal *et al*. 2007; Smith *et al*. 2011; Cleveland *et al*. 2010) using data from questionnaires or personal interviews. These studies have

used a wide variety of input variables for the LCA and, not surprisingly, have revealed a diversity of empirical subtypes, with most classes emerging based largely on the class of abused substances (e.g. (Lynskey *et al*. 2006; Agrawal *et al*. 2007; Cleveland *et al*. 2010; Schwartz *et al*. 2010) and/or class and mode of ingestion (Kuramoto *et al*. 2011).

Our group has been undertaking a detailed investigation of the epidemiology and genetic epidemiology of DA as defined from public records from the whole of Sweden (Kendler *et al*. 2012*a*, *b*). We undertook the present study to inform our future analyses regarding the potential heterogeneity within the broad syndrome of DA defined by the public sources available to us. In recent years, most genetically informed studies of DA, in addition to looking at the substance involved, have either examined only a single category (e.g. drug dependence) or at most divided their analysis up into the DSM-IV categories of abuse *versus* dependence (APA, 1994; Cadoret *et al*. 1995; Tsuang *et al*. 1996; Lynskey *et al*. 2002). We also hope that this study will contribute more broadly to our understanding of the typology of DA. To our knowledge, no prior study has used a national sample of subjects for an LCA based on the kind of objective information available from registries.

This study had three major aims. First, we performed an LCA on all individuals in Sweden born 1950–1993 who were registered with DA and/or its consequences in medical, legal, pharmacy and death registers ($n = 192\,501$). We determined the best-fitting LCA solution and describe the resulting classes. Second, we attempted to validate these classes by examining the demographic and registry data available to us. Do the identified classes differ meaningfully in their age at first registration, gender distribution, educational background, history of alcohol, criminal or psychiatric problems, or substance abused? Third, we were particularly interested in the impact of familial factors on DA. Hence our last aim was to examine whether our DA classes differed in risk for DA observed in their siblings and to determine the pattern of aggregation and co-aggregation among our DA classes within the large sample of full sibling pairs available to us. This is of particular interest to us given our focus on the role of familial/genetic factors in the etiology of DA.

Method

As in previous work (Kendler *et al*. 2012*a*, *b*), in this study we used linked data from multiple Swedish nationwide registers. Linking was achieved through the unique 10-digit personal identification number assigned at birth or immigration to all Swedish residents. Each personal identification number is replaced with a serial number to ensure anonymity of individuals. Our database contained eight sources:

(1) The Swedish Hospital Discharge Register, which includes all hospitalizations for people in Sweden from 1964 to 2009.
(2) The Swedish Prescribed Drug Register, which includes all prescriptions in Sweden picked up by patients from 1 July 2005 to 31 December 2009.
(3) The Swedish Cause of Death Register, which contains all causes of death and date of death from 1961 to 2010.
(4) The Multi-Generation Register, which includes information on family relationships for all individuals born in Sweden in 1932 onwards.
(5) The Out-patient Care Register, which includes information from out-patient clinics covering all geographic regions in Sweden from 2001 to 2009, with information on increasing number of clinics for each year during this period.
(6) The Primary Health Care Register, which includes out-patient primary care data on diagnoses and dates of diagnoses for 1 million patients from Stockholm and the middle part of Sweden (2001–2007).
(7) The Swedish Crime Register, which includes nationwide data on all convictions, including those for DA, from 1973 to 2011.
(8) The Swedish Suspicion Register, which includes nationwide data on all individuals strongly suspected of crime, including DA, from 1998 to 2011.

DA was identified in the Swedish medical registers (registers 1, 5 and 6 above) and the Cause of Death Register by ICD codes (ICD-8: Drug dependence (304); ICD-9: Drug psychoses (292) and Drug dependence (304); ICD-10: 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, which 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 (on average) more than four defined daily doses a day for 12 months of either sedatives/hypnotics [Anatomical Therapeutic Chemical (ATC) Classification System N05C and N05BA] or opioids (ATC: N02A). We restricted the diagnosis of DA to individuals above the age of 10 except for the Prescribed Drug Register,

where the age limit was set at 18 years. This study was approved by the Ethics Committee of Lund University in Malmö, Sweden.

Sample

The LCA database for these analyses was created by entering all individuals in the Swedish population who were born between 1950 and 1993 and registered with DA between 1973 and 2011. Using register information, we created seven variables defining whether each individual was registered with DA based on data in the Swedish Crime Register, the Swedish Suspicion Register, the Swedish Hospital Discharge Register, the Out-patient Care Register, the Primary Health Care Register, the Swedish Prescribed Drug Register and/or the Cause of Death Register. Based on these seven variables we created the following four dichotomous variables, as the LCA requires that all observed variables within each latent class be statistically independent: Crime (registered in the Crime Register and/or the Suspicion Register or not), Medical (registered in the Hospital Discharge Register, the Out-patient Register and/or the Primary Health Care Register or not), Prescribed Drugs (registered in the Prescribed Drug Register or not), and Mortality (registered in the Cause of Death Register or not). In addition, we summarized the number of times each individual was registered in any register. We created three groups based on this variable: individuals registered (i) once, (ii) two to three times, and (iii) four times or more. As the Crime Register and the Suspicion Register could overlap, we required that each registration in the Suspicion Register, to be counted, could not be followed by a registration in the Crime Register within a year. In total, we entered five observed items into the LCA, including three categories of numbers of registrations.

We then collected variables that could function as external validators for the LCA. We included year of birth, sex (male or female), age at first registration with DA, education, alcohol use disorders (AUDs), psychiatric disorders, and non-DA-related criminal behavior. Education was categorized into two groups: low ($\leqslant$9 years in school) and high ($>$9 years in school). For individuals born between 1983 and 1993, we used their parents' education as a proxy (selecting the parent with the higher number of years of education). AUDs in the medical registers were defined according to the following ICD codes: ICD-9: 291, 303, 305A and V79B; and ICD-10: F10 (excluding F10.0), Z50.2 and Z71.4. AUDs in the Swedish Prescribed Drug Register were identified by the following ATC codes: N07BB01, N07BB03 and N07BB04. Psychiatric disorders were defined according to the following codes in the medical registers: ICD-9: 31, 290, 293–299 and 300–309; and ICD-10: F0 and F2–F9. Criminal behavior (violent crime, property crime, sexual crime and/or gun crime) was defined according to the following laws in the Crime Register (chapter/paragraph: 3/5, 3/6, 4/5, 17/1, 17/2, 4/7, 8/5, 8/6, 6/1–6/10, 6/12, 16/11a, 4/4, 13/1, 13/2 and 3/1–3/3). ICD codes from the medical registers indicate the abused substance. No such information was available from the Crime Register. We included the following substances as additional external validators: opiates (ICD-9: 304A; ICD-10: F11), cannabis (ICD-9: 304D; ICD-10: F12), sedatives/hypnotics (ICD-9: 304W 304B, 304H and 304X; ICD-10: F13 and F19), cocaine (ICD-9: 304C; ICD-10: F14), stimulants (ICD-9: 304E; ICD-10: F15), hallucinogens (ICD-9: 304F; ICD-10: F16) and solvents (ICD-9: 304G; ICD-10: F18).

In a further attempt to validate the LCA we created a sibling database where we double entered all full siblings pairs (excluding twins) in the Swedish population where both siblings in the pair were born between 1950 and 1993. We also required that both siblings in the pair were alive after 1972, and that both siblings were alive at the age of 15. We linked this sibling database to assigned class membership data from the LCA.

Statistical methods

LCA was used to identify homogeneous DA groups based on different types of observed registrations. The number of latent classes indicated by the observed variables was determined by comparing model fit statistics between nested models. Improvement in model fit is indicated by smaller values of G^2, Akaike's Information Criterion (AIC; Akaike, 1987), the Bayesian Information Criterion (BIC; Schwarz, 1978) and entropy values (Lanza *et al*. 2007) close to 1.0. However, as the number of classes is influenced by the number of observed variables, both empirical (improved model fit) and theoretical (model interpretability) aspects were considered. Individual subjects were then assigned class membership based on the likelihood of their particular response profile.

The next step was to determine whether there were important differences across LCA classes in terms of seven external validators not entered into the LCA: year of birth, sex, education, age at first registration, AUDs, criminal behavior, and psychiatric disorders. We used χ^2 analyses for categorical variables and one-way ANOVA for continuous variables. For latent classes with high assignment probabilities from the medical registers, we also investigated, by χ^2 analyses, differences across classes regarding the abused substance. Thereafter, we examined, by

https://doi.org/10.1017/S0033291713000081 Published online by Cambridge University Press

Table 1. *Fit indices for the latent class analysis of drug abuse*

No. of latent classes	Log-likelihood	AIC	BIC	Entropy	df
2	−460 383.03	83 538.53	83 670.71	1.00	34
3	−435 779.72	34 345.92	34 549.27	0.87	27
4	−422 884.41	8 569.29	8 843.82	0.89	20
5	−420 956.90	4 728.28	5 073.99	0.87	13
6	−419 024.28	877.04	1 293.93	0.90	6

AIC, Akaike's Information Criterion; BIC, Bayesian Information Criterion; df, degrees of freedom.

Table 2. *Assignment probabilities by class*

	Class 1 Low-frequency pure criminal	Class 2 High-frequency medical criminal	Class 3 Low-frequency pure medical	Class 4 High-frequency medical	Class 5 Prescription	Class 6 Death
Class membership probabilities (S.E.)	0.4511 (0.0043)	0.2720 (0.0042)	0.1611 (0.0012)	0.0816 (0.0011)	0.0322 (0.0004)	0.0020 (0.001)
Item response probabilities						
Prescribed	0	0.0358	0	0.1601	1	0
Death	0.0002	0.0055	0	0.0490	0.0002	0.9996
Medical Register	0	0.7090	1	1	0	0.0002
Crime Register	1	0.9875	0	0.0003	0.0010	0.0003
No. of registrations						
1	0.4556	0	0.7735	0	0.4359	0.9997
2–3	0.3507	0.1253	0.2264	0.2726	0.3387	0.0001
⩾4	0.1937	0.8747	0	0.7274	0.2254	0.0002

S.E., Standard error.

logistic regression, the patterns of the external validators when comparing the four major LCA classes.

To further validate the LCA solution, we investigated, in the sibling database, patterns of concordance among siblings for the different LCA classes. Tetrachoric correlations and odds ratios (ORs) were calculated for sibling pairs across LCA classes. Statistical analyses were performed using PROC LCA in SAS v. 9.2 (Lanza *et al*. 2011).

Results

Background data

Of the 192 501 individuals identified with DA in Sweden during the period 1973–2011, the proportions with registrations from our four ascertainment sources were: prescription 5.5%, death 0.8%, medical 43.6%, and criminal 72.0%. We found that 34.6% of individuals were registered only once, 26.2% two to three times and 39.2% four or more times. The mean (S.D.) age at first registration was 28.7 (10.4) years. Men constituted 73.3% of the cases and 29.5% met our criteria for low educational status. Of the individuals identified with DA, 26.9% were also registered as having an AUD, 37.8% as having a psychiatric disorder, and 59.1% as having a criminal record (other than possible criminal registration for DA).

Results of the LCA

Using the chosen variables, we had sufficient degrees of freedom to fit up to a six-class model. All the fit indices continued to improve with increasing number of classes (Table 1), with the exception of the entropy index, which minimized at five classes. We therefore considered the six-class solution for further analysis.

The item response probabilities for these six classes are shown in Table 2. Class 1 had by far the highest class membership probability (45.1%) and almost all of the members were registered for DA through criminal records only (only a very small percentage were also registered for DA at death). We termed this class the 'low-frequency pure criminal class' because

https://doi.org/10.1017/S0033291713000081 Published online by Cambridge University Press

Table 3. *Comparison of potential validators across classes*

	Class 1 Low-frequency pure criminal	Class 2 High-frequency medical criminal	Class 3 Low-frequency pure medical	Class 4 High-frequency medical	Class 5 Prescription	Class 6 Death	χ^2/F value (p value)
Most probable class membership, *n* (%)	101 257 (53)	37 495 (19)	34 481 (18)	12 687 (7)	6 204 (3)	377 (0)	
Year of birth (mean)	1 980	1 973	1 967	1 963	1 962	1 960	12 849 (<0.0001)
Men (%)	84	77	49	57	45	86	20 761 (<0.0001)
Low education (%)	25	39	30	41	27	76	3 806 (<0.0001)
Age at first registration (years), mean (S.D.)	26 (9)	26 (8)	31 (11)	31 (10)	44 (10)	32 (8)	5 910 (<0.0001)
Alcohol use disorders (%)	13	48	37	57	18	39	26 598 (<0.0001)
Psychiatric disorders (%)	18	53	62	73	60	16	37 898 (<0.0001)
Criminal behavior (%)	60	86	37	51	23	35	22 094 (<0.0001)

S.D., Standard deviation.

more than 80% of the members of this class were registered for DA up to three times.

Class 2, with a membership probability of 27.2%, was the only class where most members were registered for DA in multiple ways, most often from criminal and medical sources. We called this the 'high-frequency medical criminal class' because all individuals in this class were registered at least twice, and more than 87% of the sample were registered at least four times.

Subjects in class 3, with a class membership probability of 16.1%, contained solely individuals registered for DA from medical sources. We called this the 'low-frequency pure medical class' because no one in this class was registered for DA more than three times.

Class 4, which had a membership probability of 8.2%, consisted of all individuals who had been registered with DA from medical sources and a relatively small proportion of them had also been registered from the prescription, death and crime registers. We called this the 'high-frequency medical class' because no members of this class had only one DA registration and 73% were registered more than four times.

Class 5, with a class membership probability of only 4.2% of the sample, contained all individuals who had been registered in the prescription registry and a very small number had also been registered in the death or crime registers. Self-evidently, this was the 'prescription class'.

Class 6 had a membership probability of only 0.2% and consisted almost entirely of individuals whose only registration for DA was by the death registry. This was called the 'death class'.

The mean posterior probabilities of class membership were: 86, 100, 90, 97, 100 and 100% in classes 1–6, respectively.

Validation of latent classes

The first row of Table 3 shows the number of individuals assigned to each class. The remainder of the table compares individuals assigned to these six classes, on seven potential validators. All of them differed highly significantly across classes. The classes varied substantially in age, with a 20-year difference in mean year of birth. The most common classes 1, 2 and 3 were substantially younger than the rarer classes 4, 5 and 6. The classes could be divided into two groups by gender composition, with classes 1, 2 and 6 having a large male excess and classes 3, 4 and 5 having nearly equal numbers of men and women. Classes 1 and 5 were the best educated, and class 4, and especially class 6, the least well educated. We saw an 18-year span across classes in mean age at first registration, with classes 1 and 2 being the youngest and classes 5 and 6 the oldest

https://doi.org/10.1017/S0033291713000081 Published online by Cambridge University Press

Table 4. *Results of validators from a multivariate logistic regression*

	Class 1 Low-frequency pure criminal (0) *v.* class 2 high-frequency medical criminal (1)	Class 3 Low-frequency pure medical (0) *v.* class 1 low-frequency pure criminal (1)	Class 3 Low-frequency pure medical (0) *v.* class 2 high-frequency medical criminal (1)	Class 4 High-frequency medical (0) *v.* class 1 low-frequency pure criminal (1)	Class 2 High-frequency medical-criminal (0) *v.* class 4 high-frequency medical (1)	Class 3 Low-frequency pure medical (0) *v.* class 4 high-frequency medical (1)
Year of birth (1 year increase, centered at 1975)	1.21 (1.22–1.22)	0.82 (0.82–0.82)	0.96 (0.96–0.97)	0.77 (0.77–0.77)	1.10 (1.10–1.11)	1.05 (1.05–1.05)
Sex (men *v.* women)	0.58 (0.56–0.60)	4.44 (4.28–4.61)	2.26 (2.17–2.35)	3.94 (3.71–4.20)	0.49 (0.47–0.52)	0.99 (0.95–1.04)
Low education	1.10 (1.06–1.14)	1.16 (1.12–1.21)	1.22 (1.17–1.27)	0.99 (0.93–1.05)	1.10 (1.05–1.16)	1.24 (1.19–1.30)
Age at first registration	0.80 (0.80–0.80)	1.17 (1.16–1.17)	0.97 (0.97–0.98)	1.22 (1.21–1.22)	0.99 (0.98–0.99)	0.97 (0.97–0.97)
AUD	2.98 (2.87–3.09)	0.50 (0.48–0.53)	1.58 (1.52–1.65)	0.43 (0.40–0.46)	0.95 (0.90–1.00)	1.52 (1.45–1.59)
Crime	1.91 (1.83–1.98)	3.36 (3.24–3.48)	8.43 (8.11–8.78)	2.76 (2.60–2.94)	0.15 (0.15–0.16)	1.62 (1.55–1.69)
Psychiatric diagnoses	3.55 (3.44–3.66)	0.19 (0.19–0.20)	0.78 (0.75–0.81)	0.11 (0.11–0.12)	2.14 (2.03–2.26)	1.62 (1.54–1.70)

AUD, Alcohol use disorders.

at registration. Rates of AUD differed markedly across classes, with classes 2 and 4 having high rates of AUDs and classes 1 and 5 having relatively low rates. Rates of registration for psychiatric illness also varied widely, with classes 1 and 6 standing out for their low rates compared to the other classes. Finally, rates of non-drug criminal registrations differed substantially across classes, being lowest in 5 and 6, and in 2 and 1.

For our four most common classes (1–4) taken two at a time (Table 4), we examined in a multivariate logistic regression the ability of these seven validators to distinguish between them. Nearly all of these comparisons were statistically significant. For example, compared to class 1, class 2 is distinguished by being significantly younger, more frequently female, more likely to have lower educational attainment, more likely to be young at first registration and having higher rates of AUD, psychiatric illness and non-DA criminal registration. An examination of Table 4 suggests that the variables that most strongly discriminated among these four DA subtypes were gender and registration for AUD, psychiatric illness and non-DA crime.

Medical diagnoses for DA typically provided information about the psychoactive substance being abused. This permitted us to examine the relationship between three of our DA classes (classes 2, 3 and 4) and the abused drug recorded for those subjects medically registered for DA (all of classes 3 and 4 and 71% of class 2) (Table 5). Compared to classes 2 and 4, class 3 has much lower levels of opiate, cannabis and non-cocaine stimulants abuse and moderately lower levels of sedative/hypnotic abuse. The patterns of abused substances were relative similar in classes 2 and 4.

Finally, we attempted to validate our DA subtypes by examining their aggregation and co-aggregation within 6 528 367 Swedish full sibling pairs and 401 902 pairs containing at least one member registered for DA. The prevalence of DA in full siblings of our six classes differed significantly ($\chi^2 = 2067.08$, df = 5, $p < 0.0001$) and was highest in siblings from classes 2 ($20.0 \pm 0.2\%$), 6 ($17.4 \pm 2.1\%$) and 1 ($15.9 \pm 0.1\%$) and lowest in siblings of classes 5 ($8.9 \pm 0.3\%$), 3 ($9.5 \pm 0.5\%$) and 4 ($13.3 \pm 0.3\%$).

Table 6 provides the tetrachoric correlation and OR for each possible combination of our classes. Of relevance, for DA as a single class, we obtained in all sibling pairs a tetrachoric correlation (S.E.) of +0.39 (0.02) and an OR of 5.94 [95% confidence interval (CI) 5.87–6.02]. We focus, in this discussion, on the tetrachoric correlation as an index of familial aggregation because it is less sensitive to changes in base rates than the OR.

Four important trends are evident in this table. First, with the exception of class 6, which was too rare to

https://doi.org/10.1017/S0033291713000081 Published online by Cambridge University Press

Table 5. *Comparison of drugs (ICD codes) across classes*

	Class 2 High-frequency medical criminal	Class 3 Low-frequency pure medical	Class 4 High-frequency medical	χ^2 (p value)
n	37 495	34 481	12 687	
F11 (304A): opiates (%)	24	9	27	3593 (<0.0001)
F12 (304D): cannabis (%)	22	8	16	2607 (<0.0001)
F13 F19 (304W 304B, 304H, 304X): sedatives and hypnotics + combinations/others (%)	62	51	66	1353 (<0.0001)
F14 (304C): cocaine (%)	2	2	1	204 (<0.0001)
F15 (304E): other stimulants (%)	28	8	21	4938 (<0.0001)
F16 (304F): hallucinogens (%)	3	1	1	383 (<0.0001)
F18 (304G): solvents (%)	1	1	1	9.1 (0.0106)

analyze usefully, we saw strong evidence for familial aggregation of each class, with tetrachoric correlations ranging from a low of +0.20 (class 3) to a high of +0.40 (class 2). The ORs were all also robust, ranging from 3.98 for class 2 to 14.68 for class 5. Second, with the exception of classes 1 and 5, all the cross-class correlations were positive, suggesting some general familial vulnerability to DA. Third, for classes 1, 2, 4 and 5 aggregation within class was consistently greater than co-aggregation across classes. For example, the correlation for class 1 across sibling pairs was +0.40 and was substantially higher than any of the correlations observed between class 1 and classes 2–6. Class 3 was the exception because the within-class correlation (+0.20) was slightly exceeded by the correlation between classes 3 and 4 (+0.22). This pattern of findings suggests some specificity in the familial factors influencing class membership. Fourth, the pattern of co-aggregation provides a rough guide to the relationship of familial risk factors across classes. For example, our low-frequency pure criminal class (class 1) has the strongest co-aggregation with the other class with frequent criminal registrations for DA (class 2 high-frequency medical criminal). However, our high-frequency medical class 4 was more strongly related to class 2 than to our low-frequency medical class 3, suggesting an important familial vulnerability to frequency of DA registrations, perhaps as an index of severity of addiction. Our low-frequency medical and criminal classes were weakly inter-related, suggesting some substantial independence for the familial factors predisposing to DA seen in criminal *versus* medical settings.

Discussion

We had three aims in this report. First, we performed an LCA of DA on a complete national sample of individuals registered for DA problems in Sweden born 1950–1993. This LCA, which used the information available in the registries, identified six classes of DA with widely varying frequencies. These classes were characterized by the mode of registration for DA and the frequency of registration.

Our second aim was to determine whether these classes were 'valid' by examining how much they differed on potentially important variables that were external to the LCA. The results suggest that these classes are meaningful and relatively distinct. Our low-frequency pure criminal DA class (class 1) was largely male and young with an earlier first DA registration, low rates of AUDs and psychiatric illness and high rates of other criminal registrations. Those in our high-frequency medical criminal class (class 2) were perhaps the most severely ill with especially elevated rates for registration for alcohol, psychiatric and criminal problems and, relative to class 3, high rates of stimulant, opiate and cannabis abuse. The low-frequency pure medical class (class 3) was notable for a nearly equal gender composition, high rates of psychiatric illness and a high proportion, relative to other substances, of sedative/hypnotic abuse. Compared to class 3, the high-frequency pure medical class had more males, lower education, higher rates of alcohol, psychiatric and criminal registration, and were much more likely to have had opiate and stimulant abuse. The prescription class (class 5) stood out by a much later age at first registration, a majority of women, low rates of AUDs and crime and relatively high rates of psychiatric illness. The death class (class 6) was, not surprisingly, the oldest but was also characterized by low educational attainment, a very high proportion of males and low overall rates for AUDs, psychiatric illness and criminal registration.

Our third aim was to examine our latent classes in full sibling pairs. We first found that the familial

https://doi.org/10.1017/S0033291713000081 Published online by Cambridge University Press

Table 6. *Tetrachoric correlations and odds ratios between different latent classes in all sibling pairs in Sweden*

	Class 1 Low-frequency pure criminal	Class 2 High-frequency medical criminal	Class 3 Low-frequency pure medical	Class 4 High-frequency medical	Class 5 Prescription	Class 6 Death
Class 1: Low-frequency pure criminal	**0.398 (0.002)**	0.283 (0.003)	0.085 (0.005)	0.061 (0.007)	−0.025 (0.012)	0.076 (0.032)
	7.87 (7.72–8.02)	5.45 (5.26–5.64)	1.77 (1.67–1.88)	1.59 (1.43–1.76)	0.80 (0.66–0.98)	2.10 (1.18–3.74)
Class 2: High-frequency medical criminal		**0.404 (0.004)**	0.205 (0.005)	0.253 (0.007)	0.155 (0.011)	0.241 (0.028)
		12.62 (12.15–13.11)	4.16 (3.91–4.43)	6.60 (6.07–7.18)	3.67 (3.14–4.29)	10.14 (6.52–15.78)
Class 3: Low-frequency pure medical			**0.195 (0.005)**	0.218 (0.007)	0.151 (0.011)	0.189 (0.033)
			3.98 (3.72–4.25)	5.35 (4.86–5.89)	3.62 (3.07–4.26)	6.66 (3.83–11.60)
Class 4: High-frequency medical				**0.321 (0.008)**	0.203 (0.014)	0.152 (0.050)
				13.84 (11.55–14.26)	6.27 (5.10–7.72)	5.66 (2.11–15.16)
Class 5: Prescription					**0.297 (0.013)**	0.285 (0.040)
					14.68 (12.08–17.84)	23.23 (11.51–46.86)
Class 6: Death						**0**
						0

Values are given as tetrachoric correlation (standard error) and odds ratio (95% confidence interval).

'severity' of these classes, as indexed by the rates of DA in their siblings, differed widely. Our high-frequency medical criminal class (class 2), which was the most clinically 'severe', was also the most familial and had a risk for DA in their siblings more than twice that seen in the siblings of our prescription class (class 5). Our high-frequency pure medical class had a risk for DA in siblings 40% higher than our low-frequency pure medical class, again suggesting a relationship between clinical and familial severity. Of interest, our low-frequency pure criminal class (class 1) had a risk for DA in siblings 67% higher than our low-frequency pure medical class (class 3), suggesting that a higher familial liability may be associated with a criminal rather than with a medical path to detection for DA.

We then examined the patterns of occurrence of all of our subtypes in sibling pairs. Of note, we found substantial evidence for specificity of familial transmission. That is, in four of our six classes, the pair resemblance was higher within the DA class than across classes. We also saw a meaningful pattern of co-aggregation, for example with evidence for a relatively strong association within sibling pairs between our two criminal and our two medical classes.

It is difficult to compare our findings with prior typological and LCA studies of DA because of the differences in the variables examined. For example, Alterman & Cacciola (1991) explored whether the presence of antisocial personality disorder defined a coherent DA subtype. As antisocial personality disorder is likely to be at least moderately correlated with our criminal subtypes, our results would be broadly supportive that this is an important dimension to consider for DA typologies. Most of the prior LCAs have largely sorted DA subjects on the basis of the substance(s) used (e.g. Wilkinson *et al*. 1987; Lynskey *et al*. 2006; Agrawal *et al*. 2007; Cleveland *et al*. 2010; Schwartz *et al*. 2010; Smith *et al*. 2011) and so are of limited relevance to our findings. Several studies have attempted to apply Babor's type A/B typology, originally developed for alcohol dependence (Babor *et al*. 1992), to drug abuse (Ball *et al*. 1995; Basu *et al*. 2004). Some resemblance is seen between type B cases (typified by early age at onset, high familial risk, severe drug and alcohol use, and frequent antisocial behavior) and our high-frequency medical criminal class.

The prior report most relevant to our current findings was a typological study by Byqvist & Olsson (1998) on 698 male Swedish drug addicts. Using both interviews and official records, they constructed a typology of DA based initially on degree of criminal involvement that was then verified by cluster analysis. They identified four types of DA: addicted criminals,

criminal addicts, low-crime addicts, and emotionally unstable addicts with little or no criminality. For the addicted criminals, their criminal careers preceded their DA. They had an early onset, a high number of offenses, low education and high rates of alcoholism. This group most closely resembles our high-frequency medical criminal class. For both the criminal addicts and the low crime addicts, their DA drove their criminal behavior. They had a later onset of problems, less deprivation, and fewer overall offenses and alcohol problems. Some of the features of these classes resemble our low-frequency pure criminal DA class. The emotionally unstable type of Byqvist & Olsson comprised the least disadvantaged group, characterized by high rates of psychiatric illness and minimal criminal behavior. This group shared features with our low-frequency pure medical, high-frequency medical and prescription classes.

Consistent with the conclusions of many clinicians and researchers, our results support the hypothesis of substantial heterogeneity within the syndrome of DA. Advances in our understanding of etiological processes will probably be aided by a consideration of this heterogeneity. In our further work with the rich data available to us in Sweden, we hope to contribute to an understanding of the risk factors that predispose to the broad DA syndrome and those that are of importance only for certain subgroups, including those identified in this report.

Limitations

These results should be interpreted in the context of six potentially important methodological limitations. First, we detected subjects with DA from medical, legal and pharmacy records. Although this method has the major advantage of not relying on cooperation or accurate respondent recall, our data probably contain both false-negative and false-positive diagnoses. However, an epidemiological study of DA conducted in neighboring Norway, which has similar rates of drug use and abuse (Kraus *et al*. 2003; Hibell *et al*. 2007), found lifetime prevalence rates of DSM-III-R (APA, 1987) DA similar to those found using our registry-based methods (Kringlen *et al*. 2001). Thus, major under-ascertainment of DA by other methods is unlikely. However, it is perhaps more accurate to say that we have studied the consequences or correlates of DA rather than DA *per se*. Second, the data that we could use to conduct and validate our LCA were limited, especially compared to those available in personal interview studies, and did not include, for example, information about personality and detailed measures of psychoactive substance use across all subjects.

Third, our methods of detecting DA changed over the ascertainment period of our study (1961–2011 for birth years 1950–1993). Three sources (Hospital Discharge, Death and Crime registers) began early in our ascertainment period (1961–1973) whereas three other sources (the Suspicion, Out-patient and Primary Health Care Registers) began later (1998–2001). We used all sources to maximize our case finding. Siblings, being of similar ages, would have been exposed to similar sources of detection if they developed DA. Therefore, the chances of biases being introduced by these changing methods of detection are small. Our multivariate analyses presented in Table 4 included year of birth and so would have corrected for any potential confounding effects of age.

Fourth, we did not formally correct for the correlations within siblings in our LCAs. We repeated a subset of our analysis with a robust variance estimator and it made minimal differences in the precision of our estimates.

Fifth, some of the subjects in the LCA are not assigned to individual classes with high confidence. We repeated the analyses contained in Table 3 including only LCA assignments of greater than 70% probability. Class membership declined modestly in classes 1 and 3, slightly in class 2 and not at all in the other classes. Despite the reduction in numbers classified, the pattern of validators observed in Table 3 hardly changed, suggesting the overall robustness of our findings.

Sixth, LCA assumes the independence of all relevant variables within classes. This assumption of 'local independence' is unlikely to be fully met in this sample.

Acknowledgments

This study was funded by grant RO1 DA030005 from the National Institute of Drug Abuse, the Swedish Research Council 2008-3110, 2008-2638, the ALF project grant, Lund, Sweden and the Swedish Council for Information on Alcohol and Other Drugs (CAN).

Declaration of Interest

None.

References

Agrawal A, Lynskey MT, Madden PA, Bucholz KK, Heath AC (2007). A latent class analysis of illicit drug abuse/dependence: results from the National Epidemiological Survey on Alcohol and Related Conditions. *Addiction* **102**, 94–104.

Akaike H (1987). Factor analysis and AIC. *Psychometrika* **52**, 317–332.

Alterman AI, Cacciola JS (1991). The antisocial personality disorder diagnosis in substance abusers: problems and issues. *Journal of Nervous and Mental Disease* **179**, 401–409.

APA (1987). *Diagnostic and Statistical Manual of Mental Disorders, Third Edition Revised*. American Psychiatric Association: Washington, DC.

APA (1994). *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*. American Psychiatric Association: Washington, DC.

Babor TF, Caetano R (2006). Subtypes of substance dependence and abuse: implications for diagnostic classification and empirical research. *Addiction* **101** (Suppl. 1), 104–110.

Babor TF, Hofmann M, DelBoca FK, Hesselbrock V, Meyer RE, Dolinsky ZS, Rounsaville B (1992). Types of alcoholics, I. Evidence for an empirically derived typology based on indicators of vulnerability and severity. *Archives of General Psychiatry* **49**, 599–608.

Ball SA, Carroll KM, Babor TF, Rounsaville BJ (1995). Subtypes of cocaine abusers: support for a type A-type B distinction. *Journal of Consulting and Clinical Psychology* **63**, 115–124.

Basu D, Ball SA, Feinn R, Gelernter J, Kranzler HR (2004). Typologies of drug dependence: comparative validity of a multivariate and four univariate models. *Drug and Alcohol Dependence* **73**, 289–300.

Byqvist S, Olsson B (1998). Male drug abuse, criminality and subcultural affiliation in a career perspective. *Journal of Psychoactive Drugs* **30**, 53–68.

Cadoret RJ, Yates WR, Troughton E, Woodworth G, Stewart MA (1995). Adoption study demonstrating two genetic pathways to drug abuse. *Archives of General Psychiatry* **52**, 42–52.

Cleveland MJ, Collins LM, Lanza ST, Greenberg MT, Feinberg ME (2010). Does individual risk moderate the effect of contextual-level protective factors? A latent class analysis of substance use. *Journal of Prevention and Intervention in the Community* **38**, 213–228.

Hibell B, Guttormsson U, Ahlstrom S, Balakireva O, Bjarnason T, Kokkevi A, Kraus L (2007). *The 2007 ESPAD Report: Substance Use Among Students in 35 European Countries*. The Swedish Council for Information on Alcohol and Other Drugs (CAN): Sweden.

Kendler KS, Sundquist K, Ohlsson H, Palmer K, Maes H, Winkleby MA, Sundquist J (2012*a*). Genetic and familial-environmental influences on risk for drug abuse: a national Swedish adoption study. *Archives of General Psychiatry* **69**, 690–697.

Kendler KS, Ohlsson H, Sundquist K, Sundquist J (2012*b*). Within-family environmental transmission of drug abuse: a Swedish national study. *Archives of General Psychiatry*. Published online: 10 December 2012. doi:10.1001/jamapsychiatry.2013.276.

Kraus L, Augustin R, Frischer M, Kummler P, Uhl A, Wiessing L (2003). Estimating prevalence of problem drug use at national level in countries of the European Union and Norway. *Addiction* **98**, 471–485.

Kringlen E, Torgersen S, Cramer V (2001). A Norwegian psychiatric epidemiological study. *American Journal of Psychiatry* **158**, 1091–1098.

Kuramoto SJ, Bohnert AS, Latkin CA (2011). Understanding subtypes of inner-city drug users with a latent class approach. *Drug and Alcohol Dependence* **118**, 237–243.

Lanza ST, Collins LM, Lemmon DR, Schafer JL (2007). PROC LCA: a SAS procedure for latent class analysis. *Structural Equation Modeling* **14**, 671–694.

Lanza ST, Dziak JJ, Huang L, Xu S, Collins LM (2011). *PROC LCA and PROC LTA Users' Guide* (*Version 1.2.7*). The Methodology Center, Penn State: University Park, PA.

Lynskey MT, Agrawal A, Bucholz KK, Nelson EC, Madden PA, Todorov AA, Grant JD, Martin NG, Heath AC (2006). Subtypes of illicit drug users: a latent class analysis of data from an Australian twin sample. *Twin Research and Human Genetics* **9**, 523–530.

Lynskey MT, Heath AC, Nelson EC, Bucholz KK, Madden PA, Slutske WS, Statham DJ, Martin NG (2002). Genetic and environmental contributions to cannabis dependence in a national young adult twin sample. *Psychological Medicine* **32**, 195–207.

Schwartz B, Wetzler S, Swanson A, Sung SC (2010). Subtyping of substance use disorders in a high-risk welfare-to-work sample: a latent class analysis. *Journal of Substance Abuse Treatment* **38**, 366–374.

Schwarz G (1978). Estimating the dimension of a model. *Annual Statistics* **6**, 461–464.

Smith GW, Farrell M, Bunting BP, Houston JE, Shevlin M (2011). Patterns of polydrug use in Great Britain: findings from a national household population survey. *Drug and Alcohol Dependence* **113**, 222–228.

Tsuang MT, Lyons MJ, Eisen SA, Goldberg J, True W, Lin N, Meyer JM, Toomey R, Faraone SV, Eaves L (1996). Genetic influences on DSM-III-R drug abuse and dependence: a study of 3,372 twin pairs. *American Journal of Medical Genetics* **67**, 473–477.

Wilkinson DA, Leigh GM, Cordingley J, Martin GW, Lei H (1987). Dimensions of multiple drug use and a typology of drug users. *British Journal of Addiction* **82**, 259–273.