EPIDEMIOLOGY

Patterns of Alcohol Use in Early Adolescence Predict Problem Use at Age 16

Jon Heron[1,*], John Macleod[1], Marcus R. Munafò[2], Roberto Melotti[1], Glyn Lewis[1], Kate Tilling[1] and Matthew Hickman[1]

[1]School of Social and Community Medicine, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol BS8 2PS, UK and [2]School of Experimental Psychology, University of Bristol, 12a Priory Road, Bristol BS8 1TU, UK
*Corresponding author: Tel.: +44-117-3314565; Fax: +44-117-9287325; E-mail: jon.heron@bristol.ac.uk

Abstract — Aims: Teenagers in the UK report some of the highest rates of alcohol use in Europe. We identify patterns of alcohol use in early adolescence and relate these to hazardous and harmful alcohol use at age 16. **Methods:** In a UK birth cohort, we analysed repeated measures of alcohol use from age 13 to 15 in a sample of 7100 adolescents. Data on drinking frequency and typical consumption when drinking were modelled separately using a pair of latent class models. Classes of alcohol-use behaviour were contrasted across a range of risk factors and then to hazardous and harmful alcohol use as assessed using the Alcohol Use Disorders Identification Test scale at age 16. **Results:** Heterogeneity in drinking frequency and consumption could each be captured with three classes corresponding to low, medium and high levels. In total, 14.2% were classified as high-frequency and 8.9% as high consumption alcohol users. Socio-demographic factors, maternal substance use and the young persons' use of tobacco and cannabis were associated with class membership. At age 16, 29% were drinking hazardously and a further 5.6% were assessed as harmful drinkers. Young people in the high drinking frequency or consumption class had a 9-fold increased risk of reporting harmful drinking at age 16. **Conclusions:** By the age of 16, a substantial proportion of teenagers in this sample were drinking at levels that could be considered hazardous or harmful for an adult. Patterns of alcohol exposure in early adolescence were strongly associated with later alcohol use. Altering drinking patterns in middle adolescence has the potential to reduce harmful use in later adolescence.

INTRODUCTION

In the UK, liver disease, primarily due to alcohol misuse, is the fifth commonest cause of mortality and, unlike many other major causes of death, is on the increase (Leon and McCambridge, 2006; Williams, 2006). The increase in liver deaths correlates with the increase in alcohol consumption (Academy of Medical Sciences, 2004), and furthermore, diagnosis of cirrhosis in young people is becoming more frequent due to persistent heavy drinking from an early age (Williams, 2006). Teenagers in the UK have one of the highest rates of substance use in Europe (Advisory Council on the Misuse of Drugs, 2006): 21% of those aged 11–15 years report having used alcohol in the past week; and an estimated 13.8% of young people aged 16–19 were alcohol 'dependent' (Office of National Statistics, 2000). Alcohol consumption and hazardous drinking have increased among adolescents and young people, e.g. average weekly alcohol consumption among women aged 16–19 nearly tripled from ~5 units in 1992 to 14 units per week in 2002 (1 unit = 8 g ethanol) (Goddard, 2007). Early and frequent alcohol use is associated with adverse trajectories leading to dependence in adulthood (Pitkanen *et al.*, 2005; Hingson *et al.*, 2006) as well as several adverse outcomes during adolescence (Fergusson and Horwood, 2000; Maughan *et al.*, 2004).

The use of substances in young people is widespread; however, the change in and development of substance use are heterogeneous in terms of its onset, progression and peak usage. This has prompted a number of longitudinal studies of adolescent alcohol use within the last 20 years (e.g. Duncan and Duncan, 1994; Schulenberg *et al.*, 1996; Tucker *et al.*, 2003). There are a number of well-established advantages of longitudinal, compared with cross sectional, data in clarifying causes, trajectories and consequences of behaviours such as alcohol use (Muthen and Muthen, 2000). Recent developments in analytic techniques have been critical to this work: in particular, latent variable modelling such as longitudinal latent class analysis (LLCA) or growth mixture modelling which allows for several classes (or subpopulations) within a population, each of which has its own trajectory (Chen and Kandel, 1995; Agrawal *et al.*, 2007; Patton *et al.*, 2007). Such models can bring clarity and understanding to often-complex multivariate data sets and produce useful summaries of the data, such as latent classes or groupings that can serve as both risk factor and outcome in further data analysis.

In this paper, we describe patterns of alcohol exposure from age 13 to 15 years in adolescents participating in a UK birth cohort. We compare the results obtained when focusing on two different alcohol measures—frequency of use and typical consumption when drinking. Frequency and consumption are two familiar measures of alcohol use, and although correlated, can be associated with distinct outcomes (Academy of Medical Sciences, 2004; Spurling and Vinson, 2005). A range of potential risk factors collected earlier in the young people's lives were then used to predict class membership. Finally, drinking patterns were related to harmful alcohol use at age 16 assessed by the 10-item AUDIT scale (Babor *et al.*, 2001).

MATERIALS AND METHODS

Participants

The sample comprised participants from the Avon Longitudinal Study of Parents and Children (ALSPAC; Golding *et al.*, 2001), an ongoing population-based cohort study in the South-West of England. More detailed information on the ALSPAC study is available at http://www.alspac.bris.ac.uk and details on how to apply to use ALSPAC data can be found at http://www.bristol.ac.uk/alspac/sci-com/collab-policy/. The primary source of data collection was via self-completed questionnaires; however, since the age of 7 years, the ALSPAC study children have been invited to the

annual clinic for a variety of hands-on assessments. All aspects of the study were reviewed and approved by the ALSPAC Law and Ethics Committee, which is registered as an Institutional Review Board. Approval was also obtained from the National Health Service Local Research Ethics Committees.

Repeated measures of alcohol use

Two ordinal measures of alcohol use were derived from questions asked via computer terminals in three clinics. The median ages at attendance were 12 years 10 months, 13 years 10 months and 15 years 5 months, which we simplify to 13, 14 and 15 years, respectively. Measures created were: (a) drinking frequency—classified into three categories as: none, occasional (has had a drink in the last 6 months but does not drink weekly), weekly use; (b) typical consumption—the number of units of alcohol the respondent consumes on a typical day when they have had a drink—classified into in five categories as: none, 1 or 2, 3 or 4, 5 or 6, 7 or more.

Alcohol problem-use outcome at age 16

In a postal questionnaire administered at age 16 (median age 16 years 7 months), the young people completed the self-reported version of the 10-item Alcohol Use Disorders Identification Test (AUDIT; Babor *et al.*, 2001). Here, we used the standard cut-offs for non-hazardous (total score <8), hazardous (total score 8–15) and harmful (total score 16+). Respondents answering 'No' to the stem question on alcohol use since the age of 15 were assigned a score of zero and retained for further analyses as members of the non-hazardous group.

Risk factors

Risk factors included: (a) demographic variables collected pre-birth, which comprised sex, housing tenure (coded as owned/mortgaged, privately rented, subsidized housing rented from council/housing-association), crowding status (coded as the ratio of number of residents to number of rooms in house), maternal educational attainment (coded as no high school qualifications, high school, beyond high school) and parity (coded as whether study child is first/second/third child or greater); (b) young person's risky behaviours collected through focus clinic at age 13 years and postal questionnaire at 11 years, which comprised cigarette use and cannabis use at 13 years (yes/no), conduct problems at 11 years (score of 0–1/2–3/4+ on the conduct problems subscale of the maternal report Strengths and Difficulties Questionnaire; Goodman and Scott, 1999) and, (c) maternal substance use in the offspring's later childhood collected via questionnaire, which comprised maternal smoking at offspring age 12 years (yes/no), maternal alcohol consumption at offspring age 12 years (evidence of bingeing and high weekly consumption derived from detailed record of beers, wines and spirits consumed in previous week), maternal cannabis use at offspring age 9 years (yes/no).

Statistical analysis

LLCA was used to extract patterns of alcohol-use behaviour from data on (a) drinking frequency, and (b) typical consumption. These data consist of different response strings, for instance non-drinkers would have the response '000', while persistent heavy drinkers would respond '444' and increasing drinkers '013'. LLCA assumes that variability in response is due to a latent (unobserved) grouping. Starting with a single class, additional classes are added until the various assessments of model fit reach an acceptable level. Akin to factor analysis, this latent categorical variable should explain the associations within the set of repeated measures such that within each latent class, respondents are homogeneous and also that the repeated measures are independent of each other (i.e. there is conditional independence).

To establish the optimal number of latent classes, we used (a) the sample-size adjusted Bayesian information criterion (BIC; Schwarz, 1978), (b) the Bootstrap Likelihood Ratio Test (BLRT; Nylund *et al.*, 2007) which assesses the improvement in model fit for each additional class, (c) entropy—a measure of the separation of the classes based on the posterior class membership probabilities (Ramaswamy *et al.*, 1993) and (d) bivariate model fit information—a test of the conditional independence assumption—using Pearson's χ^2. In all models, multiple random starts were used to help achieve the optimal maximum likelihood solution. Model fitting was carried out in Mplus version 6 (Muthén and Muthén, 2010).

Risk factor and outcome modelling

The resulting classes of alcohol use from the analyses described above were then used as a pair of categorical outcome variables. A two-stage modelling approach was followed in which class assignment probabilities for the chosen LLCA models were exported from Mplus and read into Stata version 11-MP2 (StataCorp., 2009). For each respondent, 100 random draws were made from their own set of class assignment probabilities, creating 100 data sets. Variability across these data sets would reflect the uncertainty in class assignment for each respondent. The effect of risk factors on the class membership for drinking frequency and the level of consumption was then assessed through a series of univariable multinomial regression models with results pooled across data sets using Rubin's rules (Little and Rubin, 1987). The aim of these analyses was to provide support for the derived latent classes by verifying their association with established alcohol-use risk factors. Finally, the relationship between latent classes and subsequent AUDIT score was assessed using multinomial regression. Models were adjusted for the potential confounding effects of the earlier risk factors described above.

Missing data considerations

An advantage of longitudinal models such as these is that full information maximum likelihood (FIML) estimation permits the inclusion of partially complete data. In contrast to imputation, FIML directly estimates all parameters using all available data (Enders, 2001; Enders and Bandalos, 2001) and, in the case of an unconditional mixture model, is based on the assumption that missing data are missing at random (MAR) conditional on the observed repeated measures data. Owing to weak social patterning of alcohol use in this cohort, and the minor difference in alcohol use prevalence seen between complete case and partially complete samples, we feel confident in assuming MAR.

Table 1. Measures of alcohol-use in restricted (complete case) and unrestricted (all available data) samples

	13 years		14 years		15 years	
	All available data	Complete case	All available data	Complete case	All available data	Complete case
Drinking frequency						
None	4690 (77.0%)	3190 (79.3%)	3429 (60.0%)	2494 (62.0%)	915 (17.9%)	777 (19.3%)
LT weekly	1052 (17.3%)	621 (15.4%)	1869 (32.7%)	1276 (21.7%)	3078 (60.4%)	2434 (60.5%)
Weekly	351 (5.8%)	210 (5.2%)	414 (7.3%)	251 (6.2%)	1107 (21.7%)	810 (20.1%)
Total	6093	4021	5712	4021	5100	4021
Typical consumption						
None	4791 (78.7%)	3128 (80.9%)	3482 (61.0%)	2441 (63.1%)	1666 (34.1%)	1389 (35.9%)
1 or 2 units	1046 (17.2%)	611 (15.8%)	1773 (31.1%)	1168 (30.2%)	1743 (35.7%)	1392 (36.0%)
3 or 4	148 (2.4%)	69 (1.8%)	322 (5.6%)	183 (4.7%)	716 (14.7%)	539 (13.9%)
5 or 6	59 (1.0%)	33 (0.9%)	101 (1.8%)	60 (1.6%)	400 (8.2%)	291 (7.5%)
7+	46 (0.8%)	26 (0.7%)	33 (0.6%)	15 (0.4%)	364 (7.5%)	256 (6.6%)
Total	6090	3867	5711	3867	4889	3867

A shortcoming of these methods is that they deal only with non-response among the repeated measures and hence the regression models described above will be affected by any missing data among the risks factors used. The FIML approach will permit the inclusion of covariates, and this can strengthen the MAR assumption if necessary; however, any missing data among these covariates will reduce the sample available for analysis. This is also true of the two-stage approach described above where gains made through FIML may be lost in the second stage when covariates are incorporated. Models and software for dealing with covariate non-response within a longitudinal mixture model setting are still in their infancy and previous work within the substance use literature has often adopted a relatively simplistic approach to these issues such as the use of a single imputed data set (Li *et al.*, 2001; Hix-Small *et al.*, 2004). For the current analysis, with the aim of understanding patterns of alcohol use in adolescents, we took a pragmatic approach, focusing on an FIML estimation of our maximal partially complete sample. Comparisons were made between the complete- and partially complete data set in terms of the number, the prevalence and profile of the latent classes extracted, and also on any conclusions made from the risk factor modelling. We go on to discuss the limitations of this approach including potential bias that may arise due to non-response in the covariates.

RESULTS

For the measures of drinking frequency used in the longitudinal analysis, 4021 young people provided a complete set of three responses, with a further 3079 partial responders giving a total of 7100 cases. Figures for the measure of typical alcohol consumption were similar with 3867 and 3215 cases, respectively, giving a total of 7082. Preliminary work showed evidence of an association between degree of missing data and gender as well as housing tenure, parity and maternal education with complete responders being more likely to be female, have parents who own their own home, have a mother educated beyond high school and have fewer siblings. Table 1 shows there are slightly lower rates of alcohol use within the complete case samples where all three repeated measures are available.

Latent class analysis

Between one- and three-class models were compared for drinking frequency, and up to four classes were considered for typical consumption. Model fit statistics are shown in Table 2. Insufficient degrees of freedom did not permit the fitting of a four-class model for frequency. There is good support for a three-class model with both the alcohol measures. The incomplete data bring more uncertainty reflected in poorer entropy and higher values of bivariate fit, however, the three-class models are still deemed acceptable.

Figure 1 shows the extracted patterns for drinking frequency and typical consumption for the partially observed data sets. To simplify the figures, we present the within class probabilities of (a) drinking on a weekly basis, and (b) drinking three or more units on one occasion. Complete ordinal profiles of drinking behaviour are available on request from the first author. For both drinking frequency and typical consumption, we have named the resulting latent groups as high, medium and low.

The results were as follows: for drinking frequency, 53.2% were classified as low frequency with very little drinking from 13 to 14 years and occasional drinking by the age of 15; 32.5% were classified occasional drinking by 14 years and all are drinking by 15 years with a quarter drinking on a weekly basis. Finally, 14.2% were classified as high-frequency drinkers with the majority drinking throughout the time period and almost two-thirds drinking weekly by age 15. For typical alcohol consumption, 58.8% were in the low consumption class with 1–2 units by age 15 but little before this age; 32.3% were in the medium consumption class where about half had consumed 1–2 units by age 13 and a third had had 3+ units by age 15. Finally, the high consumption class of 8.9% consumed to a higher level throughout the time period with three quarters drinking 3+ units by age 15 and a third drinking 7+ units in one sitting. The prevalence of different classes of behaviour was very similar for the smaller complete case samples.

There was reasonable concordance in the latent grouping derived from the two models with two-thirds of the young people in the 'same' category for frequency and consumption. For example, of the 7082 respondents, 45% were classified as low on both frequency and consumption, 17% as medium on both and 4.2% as high. Twelve percent were classified as low consumption/medium frequency, 8.2%

Table 2. Model fit statistics for latent class analyses

	Drinking frequency			Typical consumption			
Complete-case data							
Number of classes	1	2	3	1	2	3	4
Parameters	6	13	20	12	25	38	51
aBIC	19365.7	17911.2	17816.8	22307.2	21247.1	21158.3	21198.7
Lowest entropy value	—	0.838	0.728	—	0.884	0.79	0.727
Bivariate fit	1944.2	139.3	2.1	1881.1	223.6	37.2	12.6
BLRT	—	<0.001	<0.001	—	<0.001	<0.001	0.112
Partially observed data							
Number of classes	1	2	3	1	2	3	4
Parameters	6	13	20	12	25	38	51
aBIC	27671.5	25864	25745.8	32577.2	31248.1	31141.5	31190
Lowest entropy value	—	0.819	0.712	—	0.854	0.742	0.609
Bivariate fit	3745.9	264.6	5	3693.2	462.5	70.3	38.5
BLRT	—	<0.001	<0.001	—	<0.001	<0.001	0.116

Sample: frequency of use (complete = 4021/partial = 7100), typical consumption (complete = 3867/partial = 7082). aBIC, sample-size-adjusted Bayesian information criterion. Lower values indicate preferred models. BLRT, bootstrap likelihood ratio test. High (non-significant) values indicate no evidence of improvement in fit from model with one less class. Chi-square value for assessment of bivariate fit is 65.2 (based on 48 degrees of freedom). Values below this threshold indicate no evidence of residual association between measures. That is, measurements are independent, conditional on latent class variable. Lowest entropy value corresponds to the class for which class assignment was weakest. Entropy takes values between zero and one. Higher values of entropy are preferred.
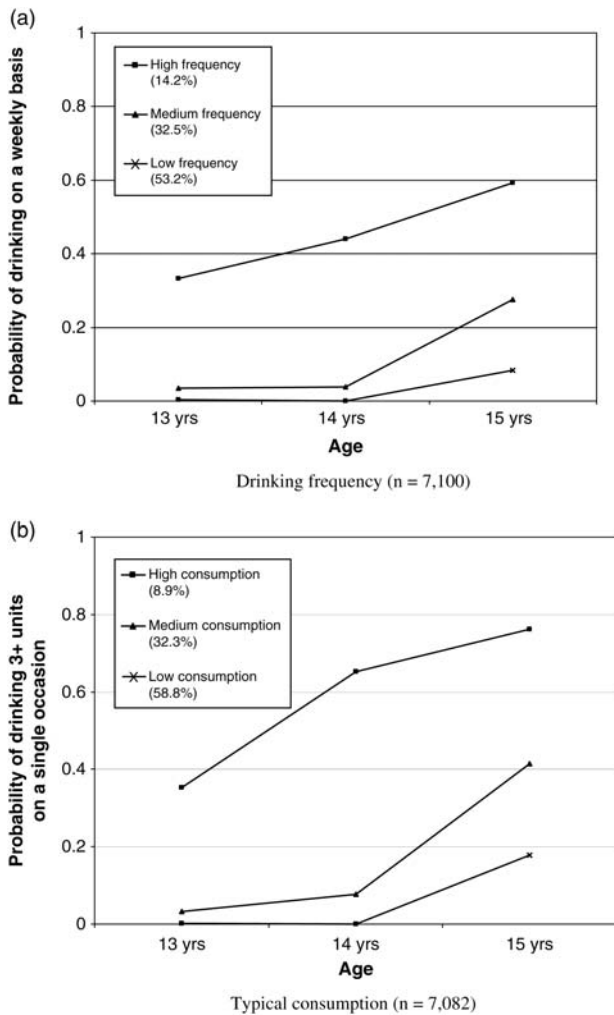

Figure 1. Resulting patterns from LLCA (partially observed data).

would be classified as medium consumption/high frequency, 7.3% as medium consumption/low frequency and 3.7% as high consumption/medium frequency class. Only 3% were classified as high/low or low/high, leaving little scope for further investigation of these highly discordant, yet potentially interesting, individuals. Prior to the assessment of risk factor associations, the resulting three-class models described above were re-estimated separately for boys and girls. There was no evidence of an improvement in fit using these models: BIC values were higher compared with the single sample model; and the patterns of response extracted by the latent class model within each gender showed good agreement between boys and girls.

Risk factors for latent class membership

Table 3 shows the relationship between adolescent alcohol-use risk factors and the latent classes for drinking frequency and consumption. These results are based on the latent classes derived using FIML estimation and the partially observed data set, and hence any drop in sample size is due to missing data in the risk factors. Conclusions remained relatively unchanged when examining results based on the complete case latent class model (data available from first author). Results are univariable, considering each risk factor in turn. In each case, the outcome has three categories with the normative category (low drinking frequency/typical consumption) treated as the reference. Following each test of association, a post-estimation comparison was made to investigate which factors were able to further distinguish between the medium- and high -use patterns (data not shown).

Gender and socio-demographic measures

There was little evidence that class membership varied for boys and girls ($P = 0.339$ for frequency and $P = 0.248$ for consumption). There was weak evidence for an association between measures of social position or deprivation and drinking frequency but slightly stronger evidence for associations with typical consumption. In the majority of cases, lower social position or higher deprivation was associated with an increase in the odds of being a high frequency or consumption alcohol user, but there was little effect for

Table 3. Risk factors for class membership (partially observed data for latent class models)

		Drinking frequency (reference = low frequency)				Level of consumption (reference = low consumption)		
	n	Medium frequency	High frequency	P-value	n	Medium consumption	High consumption	P-value
Sex	7100			0.339	7082			0.248
Male		1.00 ref				1.00 ref		
Female		1.09 (0.97, 1.24)	1.08 [0.91, 1.28]			1.09 (0.96, 1.24)	1.14 (0.93, 1.40)	
Housing tenure	6856			0.445	6839			0.013
Mortgaged/owned home		1.00 ref				1.00 ref		
Rented		0.83 (0.65, 1.07)	0.86 (0.62, 1.19)			0.83 (0.65, 1.05)	1.05 (0.72, 1.54)	
Subsidized housing		0.91 (0.73, 1.15)	1.10 (0.83, 1.46)			0.92 (0.73, 1.17)	1.58 (1.18, 2.14)	
Parity	6855			<0.001	6840			<0.001
First child		1.00 ref				1.00 ref		
Second child		1.16 (1.01, 1.34)	1.48 (1.23, 1.79)			1.23 (1.06, 1.41)	1.47 (1.16, 1.87)	
Third child or higher		1.22 (1.01, 1.47)	1.58 [1.25, 1.99]			1.22 (1.02, 1.46)	1.81 (1.38, 2.39)	
Overcrowding	6771			0.538	6754			0.059
≤1 person/room		1.00 ref				1.00 ref		
>1 person/room		0.84 (0.59, 1.21)	1.06 [0.71, 1.59]			0.82 (0.57, 1.16)	1.48 (0.97, 2.27)	
Maternal education	6811			0.337	6795			0.007
Qualifications beyond high school		1.00 ref				1.00 ref		
High school qualifications		1.03 (0.88, 1.19)	1.13 (0.93, 1.37)			1.03 (0.89, 1.20)	1.34 (1.05, 1.73)	
No high school qualifications		1.08 (0.91, 1.28)	1.26 (1.02, 1.55)			1.10 (0.93, 1.31)	1.68 (1.29, 2.21)	
Maternal weekly alcohol at 12 years	5230			0.005	5227			0.019
weekly use of <14 units		1.00 ref				1.00 ref		
14+ Units/week		1.19 (1.02, 1.40)	1.39 (1.13, 1.71)			1.19 (1.01, 1.39)	1.35 (1.04, 1.75)	
Maternal alcohol binge at 12 years[a]	5230			<0.001	5227			<0.001
No		1.00 ref				1.00 ref		
Yes		1.26 (1.08, 1.46)	1.50 (1.23, 1.84)			1.27 (1.09, 1.48)	1.47 (1.15, 1.89)	
Maternal smoking at 12 years	5424			<0.001	5421			<0.001
No		1.00 ref				1.00 ref		
Yes		1.30 (1.06, 1.60)	1.62 (1.27, 2.06)			1.29 (1.05, 1.58)	2.13 (1.60, 2.83)	
Maternal cannabis use at 9 years	5770			0.005	5763			0.002
No		1.00 ref				1.00 ref		
Yes		1.43 (1.00, 2.05)	1.92 (1.32, 2.79)			1.31 (0.94, 1.82)	2.23 (1.45, 3.45)	
Self-report smoking at 13 years	6317			<0.001	6316			<0.001
No		1.00 ref				1.00 ref		
Yes		3.33 (2.51, 4.44)	11.5 (8.75, 15.1)			3.78 (2.92, 4.91)	14.9 (11.1, 19.9)	
Self-report cannabis at 13 years	6338			<0.001	6337			<0.001
No		1.00 ref				1.00 ref		
Yes		7.19 (2.53, 20.4)	29.4 (11.4, 75.9)			7.57 (2.90, 19.8)	33.7 (13.5, 84.2)	
SDQ conduct problems at 11 years	5669			0.003				<0.001
Low		1.00 ref				1.00 ref		
Med		1.09 (0.93, 1.27)	1.17 (0.94, 1.46)			1.10 (0.95, 1.30)	1.34 (1.03, 1.74)	
High		1.25 (0.93, 1.67)	1.93 (1.41, 2.64)			1.28 (0.97, 1.68)	2.24 (1.52, 3.28)	

SDQ, strengths and difficulties questionnaire. All ages refer to the age of the young person when data was collected.
[a]Binge defined as 4+ units of alcohol on one occasion. One unit of alcohol is equivalent to 0.8 g ethanol.

medium-level alcohol users. For instance, having no maternal educational qualifications was associated with a 26% increased odds of being in the high-frequency class and 68% increased odds of being in the high-consumption class. Subsidized housing was also associated with the high-consumption class; however, there was no evidence of an association between housing tenure and drinking frequency. The strongest associations within these measures were for parity with an apparent dose–response relationship between the number of siblings and rates of high consumption.

Maternal substance use

There was strong evidence for associations between all three maternal substance use measures and drinking frequency and consumption class membership. Maternal alcohol consumption demonstrates a weak gradient effect. Associations were stronger for tobacco and cannabis, particularly the latter with double the odds of being in the high-drinking frequency and high-consumption classes for young people with cannabis, using mothers at age 9.

Young person characteristics

The cigarette/cannabis use at age 13 is strongly related to alcohol-use pattern, particularly for cannabis which was associated with being both a medium- and high-level alcohol user. Finally, higher levels of maternally reported conduct problems at age 11 are strongly related with high-frequency/-consumption with an approximate doubling of the odds, but associated with little additional risk for more moderate drinking levels.

Post-estimation differences between medium and high alcohol use

For many risk factors, associations were slightly stronger for the level of consumption compared with frequency of use. This is probably due to differences in the high-category prevalences (9 vs 14%) resulting in a more extreme group of

Table 4. Predicting harmful/hazardous alcohol use at age 16 using patterns of alcohol frequency and consumption

	AUDIT score					
	0–7 points non-hazardous use (65.3%)		8–15 points hazardous use (29.1%)		16 or more points harmful use (5.6%)	
	Prevalence	Unadjusted OR (95% CI)	Prevalence	Unadjusted OR (95% CI)	Prevalence	Unadjusted OR (95% CI)
Drinking frequency (*n* = 4092)						
Low (56.0%)	76.0%	1.00 ref	21.2%	1.00 ref	2.8%	1.00 ref
Medium (31.4%)	56.1%	1.00 ref	36.7%	2.34 (1.93, 2.84)	7.2%	3.48 (2.28, 5.31)
High (12.7%)	40.2%	1.00 ref	45.2%	4.03 (3.15, 5.17)	14.5%	9.87 (6.52, 14.9)
Typical consumption (*n* = 4086)						
Low (61.7%)	74.2%	1.00 ref	22.5%	1.00 ref	3.3%	1.00 ref
Medium (31.1%)	53.9%	1.00 ref	38.0%	2.33 (1.94, 2.80)	8.1%	3.32 (2.33, 4.73)
High (7.3%)	38.0%	1.00 ref	47.2%	4.11 (2.95, 5.72)	14.7%	8.63 (5.32, 14.0)

Prevalence indicates the prevalence of each AUDIT outcome group within each latent class, e.g. 76% of the low-frequency drinkers obtain an AUDIT score of 0–7 at age 16 years. Odds ratios from multinomial regression models using the non-hazardous use outcome level as the reference. Frequency and consumption measures were obtained from LLCA using partially observed data set.

users in the high category for the 'level of consumption', hence establishing a stronger dose response. Consequently, there was a greater tendency to observe differences between the medium and high use for this outcome. Socio-demographic factors conferred a mildly increased odds of being a high-level user compared with medium-level user, particularly parity (both outcomes) and subsidized housing/no education (consumption outcome only). In addition, maternal smoking and cannabis were associated with being a high-consumption user compared with medium-consumption user [smoking: odds ratio (OR) = 1.65 (1.20, 2.27), and cannabis: OR = 1.71 (1.05, 2.79)]. Finally, the young person's own characteristics were most able to distinguish between all three levels of the outcome, conferring between three and four times greater odds of being a high user for the smoking and cannabis predictors.

Relationship with harmful/hazardous alcohol use at 16

These analyses are based on the subsample of ~4100 with an estimated latent class membership and also a 16-year AUDIT outcome. At age 16, 29% of the young people were drinking hazardously and a further 5.6% were assessed as harmful drinkers. There was moderate evidence ($P = 0.017$) of an association between the gender and AUDIT score: 29.7% of girls scored 8–15 on the AUDIT scale (hazardous use) compared with 28.2% of boys, and 6.4% of girls scored 16 or higher (harmful use), compared with 4.6% of the boys. Table 4 shows evidence of a strong association between patterns of alcohol use across ages 13 and 15 and AUDIT scores at age 16. Being in the high use class for either drinking frequency or typical consumption is associated with an 8- to 10-fold increase in odds of harmful alcohol use at 16. Adjustment for gender, demographics and maternal substance use led to very minor attenuation (<5%) of the estimates displayed in Table 4.

DISCUSSION

In a large birth cohort, we use LLCA to derive patterns of alcohol frequency and consumption in 13- to 15-year-old adolescents and measured hazardous alcohol use at 16. Variability in frequency of drinking alcohol could be

captured by three latent classes with 53.2% classed as low, 32.5% as medium and 14.2% as high-frequency users within the partially observed data set of ~7000 cases. Among the high-frequency users, weekly consumption is common by the age of 13 and is practiced by the majority by 15. The larger medium-frequency group start off slower than their high-frequency peers but by age 15 a large proportion are drinking on a weekly basis. Variability in typical alcohol consumption levels could also be captured using three latent classes: with 58.8% low consumption users; 32.3% medium-consumption drinkers and 8.9% high-consumption class who were likely to be drinking three or more units by the age of 13 with two-thirds of this group drinking 3+ units on one occasion by age 15. We have shown consistent clear univariable associations between early life risk factors and latent class membership with stronger effects for the higher alcohol use/consumption groups. At age 16, over one in three of adolescents were categorized as hazardous or harmful drinkers and 16-year alcohol-use status was associated strongly with alcohol-use profile between 13 and 15 years.

Strengths and limitations

While the size of the cohort is a key strength of this study, there has been considerable dropout since the study's conception, and this impacts on analyses both when using the complete case and the partially observed data sets of alcohol measures. Nonetheless, our estimates of adolescent drinking are consistent with relevant UK school and community surveys. For example, in 2008, by age 15, 80% of young people reported drinking alcohol; 13% were weekly drinkers and average consumption (excluding non-drinkers) in the last week before the survey was ~15.5 units (NHS Information Centre, 2009). Furthermore, conclusions in terms of risk factors and the AUDIT outcome were consistent across the models utilizing complete and partially observed alcohol use data, and the profiles and class prevalences from the latent class models were also relatively robust to this dropout. Previous work examining patterns of adolescent tobacco use (Heron *et al.*, 2011) showed marked difference in the prevalence of high-consumption classes obtained with complete- and partially observed data set with lower cigarette consumption among those providing complete data. This is consistent

with the earlier finding that compared with smoking; the alcohol use in young people within the UK is more moderately and differentially socially patterned (Macleod *et al.*, 2008; Melotti *et al.*, 2011) and that social patterning across cohorts is not a consistent finding (Wiles *et al.*, 2007). There remains the issue of generalizability to the broader ALSPAC cohort as these models were based on approximately half of the cohort who initially enrolled. In this instance, we feel that the results shown here support the notion that whilst the relative sizes of the latent classes may differ slightly within the larger sample of those who originally enrolled in ALSPAC, the associations reported would be robust to this loss to follow-up.

A key limitation of many surveys of alcohol use is that the data are based on self-report. However, unlike cannabis and tobacco use, there currently exists no reliable biological confirmation of drinking behaviour so the validity of our measures may be questioned. In addition, it has been reported that the data collection context can impact on the measurement of drug-use data (Percy *et al.*, 2005; Grucza *et al.*, 2007) which raised additional issues when measuring the substance use behaviour on a cohort of this age. However, we feel confident that the approach used by ALSPAC in which cohort members attend the ALSPAC premises individually and answer questions at a computer terminal, and are given clear assurances of anonymity and confidentiality for both clinic and postal questionnaires, will minimize any tendency to exaggerate or under-estimate the alcohol use. While we cannot rule out misclassification of alcohol consumption, the reported use of alcohol appears to be consistent with other surveys in this age group.

A further limitation to the interpretation of our findings is the use of adult cut-points for the AUDIT scale. In adults, the AUDIT can also be used to categorize alcohol 'dependence' (total score 20+) but it is not certain that this is applicable to adolescents (Caetano and Babor, 2006) and hence we chose to combine these respondents with the harmful-use group scoring 16 or more. Furthermore, some studies recommend using lower cut-off scores to distinguish hazardous and harmful use (Chung *et al.*, 2000; Knight *et al.*, 2003; Cook *et al.*, 2005; Santis *et al.*, 2009), but there is no clear consensus. Without the data to derive adolescent cut-points of our own, we opted for the well-established adult grouping—which may underestimate the level of hazardous drinking.

Relating findings to current evidence

The finding of extensive alcohol use in mid-adolescents is consistent with recent UK survey data. For instance, these report that most young people by age 16 have tried alcohol by age 16, that there seem to be little difference in alcohol across the ages for girls and boys and, by age 15 ~40% of children will be drinking weekly (Currie *et al.*, 2008; NHS Information Centre, 2009). What these cross-sectional surveys, however, fail to capture is how drinking patterns evolve for different subgroups of children. While it is true that alcohol consumption increases during adolescence— there are different patterns of use—with some children engaged in high-frequency and high-consumption drinking. It is notable that our cohort and other recent surveys fail to detect differences between girls and boys due most likely to increase in consumption by girls. For example, the recent

report of the Adult Psychiatric Morbidity Survey (NHS Information Centre, 2007) reported comparative increases in alcohol dependence in young women relative to men aged 16–24.

The categorization of adolescent alcohol use as high, medium and low in terms of both frequency and quantity of consumption may seem self-evident and the added value of using latent class analysis to derive these categories may not be immediately apparent. Nevertheless, there are a number of advantages of this data-driven mixture-modelling approach. The first is the ability to deal with a large number of different response patterns—there are 125 of these for typical consumption in the complete-case data set alone; secondly this approach allows for the inclusion of more participants as a partial set of responses is permitted—it would not be trivial to assign respondents to groups by hand if their response pattern was incomplete; thirdly, by modelling latent class assignment in a probabilistic way, uncertainty can be preserved allowing all respondents to be assigned irrespective of how unusual their response pattern may be.

Previous work describing patterns of drinking in adolescents has used a variety of measures including frequency of heavy use (Tucker *et al.*, 2003; Jackson *et al.*, 2005), frequency of any use (Blozis *et al.*, 2007), recent quantity of ethanol consumed (Wiesner *et al.*, 2007) or composite measures derived from a series of questions on alcohol use and misuse (e.g. Komro *et al.*, 2010). This work has typically led to the extraction of between two and five classes, e.g. two (Li *et al.*, 2002), three (Flory *et al.*, 2004), four (Jackson *et al.*, 2008) and five (Tucker *et al.*, 2003). It has recently been argued that the number of classes identified in such analyses may be sensitive to a various aspects of the data and chosen model (Jackson and Sher, 2005, 2006, 2008) and that an apparent phenotypic variation within and between study samples may partly reflect differences in the number and type of measures used and the age-range modelled rather than true differences in substance use behaviour between cohorts. As far as we are aware, the current study is the first in the UK to apply these methods to adolescent alcohol use. We have found that changing patterns of drinking frequency and typical consumption between 13 and 15 can each be adequately summarized by three classes of behaviour. While the class prevalence differed depending on the repeated measure employed, the association between class membership and key risk factors for alcohol use was robust to the variables chosen as was the shape and interpretation of the patterns of use extracted. In a their paper which employed similar measures of drinking frequency and consumption, Casswell *et al.* (2002) concluded that quantity and frequency of drinking in adolescence have different trajectories; however, our own work does not fully support this finding. A possible explanation would be that their models spanned the ages of 18–26 years where a wider range of drinking patterns would be expected.

Implications

We found very high associations between medium and high drinking patterns and hazardous/harmful alcohol consumption at age 16, which we believe is not tautological. First, the exposure derived from data covering the period 13–15 years is not measuring hazardous alcohol use but simply indicating

medium to high levels in terms of the proportion of adolescents by age that are drinking on a weekly basis or typically consuming more than 3 units when drinking—neither of which *per se* would classify the young person as drinking hazardously. Second, our patterns span over 3 years (and cannot be distilled or adequately summarized by a single/or earliest data point) so it is not a simple case of identifying young people at a single point in time who may develop problems later. Rather the study points to the importance of reducing the average levels of drinking among young adolescents in order to reduce hazardous drinking at age 16—as young adolescents in the high classes of alcohol consumption and frequency were 9–10 times more likely to be hazardous drinkers later in adolescence.

A recent review of reviews (Newbury-Birch *et al.*, 2008) identified 'many adverse consequences of drinking alcohol during childhood and adolescence which would seem to outweigh the modest number of positive impacts'. In addition, for the first time, The Chief Medical Officer recommended that 'an alcohol-free childhood is the healthiest and best option'; that 'if children drink alcohol, it should not be until at least the age of 15 years'; 'should always be with the guidance of a parent or carer or in a supervised environment' and 'If 15–17 year olds do consume alcohol they should do so infrequently and certainly on no more than one day a week' and 'never exceed recommended adult daily limits' (Chief Medical Officer, 2009). In this context, our evidence suggests that the current alcohol use among adolescents is widespread and some distance from public health policy recommendations (with many young people drinking at an early age, and by late adolescence a significant minority are exceeding recommended limits of drinking for adults). Multi-component interventions are required in order to minimize and reduce average levels of drinking among young and mid-adolescents.

REFERENCES

Academy of Medical Sciences. (2004) *Calling Time, The Nation's Drinking as a Major Health Issue*. http://www.acmedsci.ac.uk/p48prid16.html.

Advisory Council on the Misuse of Drugs. (2006) *Pathways to Problems*. London: Home Office.

Agrawal A, Lynskey MT, Madden PAF *et al.* (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.

Babor T, Higgins-Biddle J, Saunders J *et al.* (2001) *The Alcohol Use Disorders Identification Test, Guidelines for Use in Primary Care*, 2nd edn. Geneva: World Health Organization.

Blozis SA, Feldman B, Conger RD. (2007) Adolescent alcohol use and adult alcohol disorders: a two-part random-effects model with diagnostic outcomes. *Drug Alcohol Depend* **88**(Suppl 1): S85–96.

Caetano R, Babor TF. (2006) Diagnosis of alcohol dependence in epidemiological surveys: an epidemic of youthful alcohol dependence or a case of measurement error? *Addiction* **101**(Suppl 1):111–4.

Casswell S, Pledger M, Pratap S. (2002) Trajectories of drinking from 18 to 26 years: identification and prediction. *Addiction* **97**:1427–37.

Chen K, Kandel DB. (1995) The natural history of drug use from adolescence to the mid-thirties in a general population sample. *Am J Public Health* **85**:41–7.

Chief Medical Officer. (2009) *Draft Guidance on the Consumption of Alcohol by Children and Young People*. http://webarchive.nationalarchives.gov.uk/20100202100434/http://www.dcsf.gov.uk/consultations/downloadableDocs/CMO%20Guidance.pdf.

Chung T, Colby SM, Barnett NP *et al.* (2000) Screening adolescents for problem drinking: performance of brief screens against DSM-IV alcohol diagnoses. *J Stud Alcohol* **61**:579–87.

Cook RL, Chung T, Kelly TM *et al.* (2005) Alcohol screening in young persons attending a sexually transmitted disease clinic. Comparison of AUDIT, CRAFFT, and CAGE instruments. *J Gen Intern Med* **20**:1–6.

Currie C, Gabhainn S, Godeau E. (2008) Inequalities in young people's health: international report from the HBSC 2005/06 survey. WHO Policy Series: Health policy for children and adolescents Issue 5. Copenhagen: WHO Regional Office for Europe.

Duncan SC, Duncan TE. (1994) Modeling incomplete longitudinal substance use data using latent variable growth curve methodology. *Multivariate Behav Res* **29**:313–38.

Enders CK. (2001) The performance of the full information maximum likelihood estimator in multiple regression models with missing data. *Educ Psychol Meas* **61**:713–40.

Enders CK, Bandalos DL. (2001) The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Struct Equ Model* **8**:430–57.

Fergusson DM, Horwood LJ. (2000) Alcohol abuse and crime: a fixed-effects regression analysis. *Addiction* **95**:1525–36.

Flory K, Lynam D, Milich R *et al.* (2004) Early adolescent through young adult alcohol and marijuana use trajectories: early predictors, young adult outcomes, and predictive utility. *Dev Psychopathol* **16**:193–213.

Goddard E. (2007) *Estimating Alcohol Consumption from Survey Data: Updated Method of Converting Volumes to Units*. London: ONS HMSO, No. 37.

Golding J, Pembrey M, Jones R. (2001) ALSPAC—the Avon longitudinal study of parents and children. I. Study methodology. *Paediatr Perinat Epidemiol* **15**:74–87.

Goodman R, Scott S. (1999) Comparing the Strengths and Difficulties Questionnaire and the Child Behavior Checklist: is small beautiful? *J Abnorm Child Psychol* **27**:17–24.

Grucza RA, Abbacchi AM, Przybeck TR *et al.* (2007) Discrepancies in estimates of prevalence and correlates of substance use and disorders between two national surveys. *Addiction* **102**:623–9.

Heron J, Hickman M, Macleod J *et al.* (2011) Characterizing patterns of smoking initiation in adolescence: comparison of methods for dealing with missing data. *Nicotine Tob Res* **13**:1266–75.

Hingson RW, Heeren T, Winter MR. (2006) Age at drinking onset and alcohol dependence: age at onset, duration, and severity. *Arch Pediatr Adolesc Med* **160**:739–46.

Hix-Small H, Duncan TE, Duncan SC *et al.* (2004) A multivariate associative finite growth mixture modeling approach examining adolescent alcohol and marijuana use. *J Psychopathol Behav* **26**:255–70.

Jackson KM, Sher KJ. (2005) Similarities and differences of longitudinal phenotypes across alternate indices of alcohol involvement: a methodologic comparison of trajectory approaches. *Psychol Addict Behav* **19**:339–51.

Jackson KM, Sher KJ. (2006) Comparison of longitudinal phenotypes based on number and timing of assessments: a systematic comparison of trajectory approaches II. *Psychol Addict Behav* **20**:373–84.

Jackson KM, Sher KJ. (2008) Comparison of longitudinal pheno-types based on alternate heavy drinking cut scores: a systematic comparison of trajectory approaches III. *Psychol Addict Behav* **22**:198–209.

Jackson KM, Sher KJ, Schulenberg JE. (2005) Conjoint develop-mental trajectories of young adult alcohol and tobacco use. *J Abnorm Psychol* **114**:612–26.

Jackson KM, Sher KJ, Schulenberg JE. (2008) Conjoint develop-mental trajectories of young adult substance use. *Alcohol Clin Exp Res* **32**:723–37.

Knight JR, Sherritt L, Harris SK *et al*. (2003) Validity of brief alcohol screening tests among adolescents: a comparison of the AUDIT, POSIT, CAGE, and CRAFFT. *Alcohol Clin Exp Res* **27**:67–73.

Komro KA, Tobler AL, Maldonado-Molina MM *et al*. (2010) Effects of alcohol use initiation patterns on high-risk behaviors among urban, low-income, young adolescents. *Prev Sci* **11**:14–23.

Leon DA, McCambridge J. (2006) Liver cirrhosis mortality rates in Britain from 1950 to 2002: an analysis of routine data. *Lancet* **367**:52–6.

Li F, Barrera M, Jr, Hops H *et al*. (2002) The longitudinal influence of peers on the development of alcohol use in late adolescence: a growth mixture analysis. *J Behav Med* **25**:293–315.

Li F, Duncan TE, Hops H. (2001) Examining developmental trajec-tories in adolescent alcohol use using piecewise growth mixture modeling analysis. *J Stud Alcohol* **62**:199–210.

Little RJA, Rubin DB. (1987) *Statistical Analysis with Missing Data*. New York: John Wiley.

Macleod J, Hickman M, Bowen E *et al*. (2008) Parental drug use, early adversities, later childhood problems and children's use of tobacco and alcohol at age 10: birth cohort study. *Addiction* **103**:1731–43.

Maughan B, Rowe R, Messer J *et al*. (2004) Conduct disorder and oppositional defiant disorder in a national sample: developmen-tal epidemiology. *J Child Psychol Psyc* **45**:609–21.

Melotti R, Heron J, Hickman M *et al*. (2011) Adolescent alcohol and tobacco use and early socioeconomic position: the ALSPAC birth cohort. *Pediatrics* **127**:e948–55.

Muthen B, Muthen LK. (2000) Integrating person-centered and variable-centered analyses: growth mixture modeling with latent trajectory classes. *Alcohol Clin Exp Res* **24**:882–91.

Muthén LK, Muthén BO. (2010) *Mplus User's Guide*. Sixth Edition. Los Angeles, CA: Muthén & Muthén.

Newbury-Birch D, Gilvarry E, McArdle P *et al*. (2008) *Impact of Alcohol Consumption on Young People: A Review of Reviews*. http://www.education.gov.uk/consultations/downloadableDocs/Review%20of%20existing%20reviews%20(Full).pdf.

NHS Information Centre. (2007) *Adult Psychiatric Morbidity in England 2007. Results of a household survey*.http://www.ic.nhs. uk/webfiles/publications/mental%20health/other%20mental%20health%20publications/Adult%20psychiatric%20morbidity%2007/APMS%2007%20%28FINAL%29%20Standard.pdf.

NHS Information Centre. (2009) *Smoking, Drinking and Drug Use among Young People in England in 2008*. http:www.ic.nhs.uk/pubs/sdd08fullreport.

Nylund KL, Asparouhov T, Muthén BO. (2007) Deciding on the number of classes in latent class analysis and growth mixture modelling: a Monte carlo simulation study. *Struct Equ Model* **14**:535–69.

Office of National Statistics. (2000) *ONS Psychiatric Morbidity among Adults Living in Private Households*. http://www.statistics.gov.uk/downloads/theme_health/psychmorb.pdf.

Patton GC, Coffey C, Lynskey MT *et al*. (2007) Trajectories of ado-lescent alcohol and cannabis use into young adulthood. *Addiction* **102**:607–15.

Percy A, McAlister S, Higgins K *et al*. (2005) Response consist-ency in young adolescents' drug use self-reports: a recanting rate analysis. *Addiction* **100**:189–96.

Pitkanen T, Lyyra AL, Pulkkinen L. (2005) Age of onset of drink-ing and the use of alcohol in adulthood: a follow-up study from age 8–42 for females and males. *Addiction* **100**:652–61.

Ramaswamy V, DeSabro W, Robinson W. (1993) An empirical pooling approach for estimating marketing mix elasticities with PIMS data. *Marketing Sci* **12**:103–24.

Santis R, Garmendia ML, Acuna G *et al*. (2009) The Alcohol Use Disorders Identification Test (AUDIT) as a screening instrument for adolescents. *Drug Alcohol Depend* **103**:155–8.

Schulenberg J, O'Malley PM, Bachman JG *et al*. (1996) Getting drunk and growing up: trajectories of frequent binge drinking during the transition to young adulthood. *J Stud Alcohol* **57**:289–304.

Schwarz G. (1978) Estimating the dimension of a model. *Ann Statist* **6**:461–4.

Spurling MC, Vinson DC. (2005) Alcohol-related injuries: evidence for the prevention paradox. *Ann Fam Med* **3**:47–52.

StataCorp. (2009) *Stata Statistical Software: Release 11*. College Station, TX: StataCorp LP.

Tucker JS, Orlando M, Ellickson PL. (2003) Patterns and correlates of binge drinking trajectories from early adolescence to young adulthood. *Health Psychol* **22**:79–87.

Wiesner M, Weichold K, Silbereisen RK. (2007) Trajectories of alcohol use among adolescent boys and girls: identification, val-idation, and sociodemographic characteristics. *Psychol Addict Behav* **21**:62–75.

Wiles NJ, Lingford-Hughes A, Daniel J *et al*. (2007) Socio-economic status in childhood and later alcohol use: a sys-tematic review. *Addiction* **102**:1546–63.

Williams R. (2006) Global challenges in liver disease. *Hepatology* **44**:521–6.