The clustering of risk behaviours in adolescence and health consequences in middle age

Abstract

Introduction

There is increasing interest in the clustering of risk behaviours in adolescence. However, few studies have examined what clusters of risk behaviours exist among adolescents, their early-life predictors, and their associations with later health.

Methods

We analysed data derived from 8754 participants (women 53.3%) in the 1970 British Cohort Study. Latent class analysis was used to identify clusters of risk behaviours at age 16. Regression modelling was then used to examine predictors of clusters and their consequences of risk behaviours and health outcomes at age 42.

Results

We identified two latent classes: a risky-behaviour (men: 20.0%, women: 23.6%) and less-risky-behaviour class. Among men, those in the risky-behaviour class were more likely to report smoking, multiple binge drinking, sexual debut before 16, involvement in fights and delinquency than were women. Membership in risky-behaviour class was mainly predicted by sociodemographic and parental risk behaviours and monitoring. The risky-behaviour class at age 16 was associated with the following outcome age 42: smoking status (more strongly among women), excessive alcohol consumption (more strongly among men), worse self-rated health (more strongly among men), and psychological distress (only among women).

Conclusions

Engagement in multiple risk behaviours in adolescence is an important driver of health inequalities later in life. Early life intervention, for example via school-based interventions, may be warranted for favourable lifelong health.

Keywords

Adolescent behaviours; substance abuse; life-course perspective; prospective cohort study

1. Introduction

Adolescence is a critical period for emerging risk behaviours with consequences for adult health (Patton et al., 2016). Behaviours such as smoking, drinking alcohol, drug use and sexual-risk behaviours tend to cluster within individuals (Hair, Park, Ling, & Moore, 2009; C. Jackson, Sweeting, & Haw, 2012; Lippke, Nigg, & Maddock, 2012; Meader et al., 2016). This may be because some behaviours, such as alcohol use, increase the risk of others, such as sexual-risk behaviours and cannabis use, via disinhibition (Bellis et al., 2008) or 'gateway' mechanisms (Hale & Viner, 2016; Wagner & Anthony, 2002;), or because common psychological (C. A. Jackson, Henderson, Frank, & Haw, 2012) or social factors (Alamian & Paradis, 2012) influence diverse risks. Engagement in multiple risk behaviours is of concern because the resultant harms may be summative or multiplicative (Meng, Maskarinec, Lee, & Kolonel, 1999). For example, the effects of tobacco and alcohol on some cancers are synergistic (Prabhu, Obi, & Rubenstein, 2014) as are the risks posed by unhealthy diets, excessive alcohol consumption and physical inactivity (Hart, Morrison, Batty, Mitchell, & Davey Smith, 2010; Qin, Knol, Corpeleijn, & Stolk, 2010).

A number of studies have explored groupings of adolescent engagement in multiple risks (Connell, Gilreath, & Hansen, 2009; Daw, Margolis, & Wright, 2017; Fergusson, Horwood, & Lynskey, 1994; McAloney, 2015; Pulvers et al., 2018; Sanchez-Oliva et al., 2018; Tomczyk, Isensee, & Hanewinkel, 2016; Yu, Putnick, Hendricks, & Bornstein, 2017). These often focus on a narrow range of behaviours, such as use of different substances (Tomczyk et al., 2016), clustering of risks in the same domain (e.g. physical activity (Sanchez-Oliva et al., 2018)) or clustering of substance use with sexual-risk behaviours (Connell et al., 2009; McAloney, 2015; Yu et al., 2017). Such studies tend to focus on risks defined in terms of premature engagement in adulthood behaviours such as sexual activity (Tomczyk et al., 2016) and/or engagement in

anti-social behaviours such as violence (Fergusson et al., 1994), with few such studies also encompassing other forms of risk such as physical inactivity (Pulvers et al., 2018).

Some studies examined correlates of risky-behaviour clusters, such as family structure or parental monitoring (DuRant, Smith, Kreiter, & Krowchuk, 1999; Pulvers et al., 2018). There are some studies which have documented the influence of socioeconomic status (SES) with some but not all (C. Jackson et al., 2012) reporting associations (Daw et al., 2017; Meader et al., 2016). Few longitudinal studies, however, investigated how risk factors in childhood were associated with adolescent risks (Tomczyk et al., 2016).

Some previous literature explored how involvement in risk behaviours in adolescence influenced adolescent (El Ansari, Ssewanyana, & Stock, 2018) or adult health (Champion et al., 2018). A cross-sectional study of UK university students examined groupings of smoking, physical inactivity, alcohol consumption, drug use, unhealthy nutrition and inadequate sleep (El Ansari et al., 2018). This found that engagement in multiple risks was associated with worse self-rated health and lower educational performance, particularly among men (El Ansari et al., 2018). Few studies examined associations of adolescent risk behaviours with later health: an Australian study of adolescents with 5 years follow-up reported that adolescents engaged in multiple risks experienced worse mental health in early adulthood (Champion et al., 2018). Existing studies, however, rely on relatively short follow-up periods, offering limited insights into effects in later life.

We aimed to address some of these limitations by examining clustering of risk behaviours relating not only to premature engagement in adulthood behaviours (smoking, alcohol use, sexual debut) and engagement in anti-social behaviours (drug use, fighting, delinquency, unprotected sex) but also to physical inactivity. We explored the role not only of SES and family factors but also early life and family processes as risk factors. We also for the first time examined the consequences of adolescent clustered risk for health in middle age,

examining how such associations were moderated by sex and SES so that we could assess the role of multiple risk behaviours in health inequalities. Our research questions are as follows:

1. What latent classes of clustered risk (smoking, alcohol, drug use, sexual debut before 16, unprotected sex, physical inactivity, fights and delinquency) existed among British adolescents born in 1970?
2. How is membership in latent classes predicted by family structure and functioning, and parental socioeconomic status as well as individual factors?
3. How does membership of the various latent classes for risk in adolescence predict risk and morbidity in middle age, and is this moderated by sex and socioeconomic status?

We draw on a birth cohort of individuals born in 1970, focusing on the UK because of the rich birth cohort data available. The 1970 British Cohort Study (BCS70) is particularly interesting in comprising individuals whose adolescence involved high prevalence of many risk behaviours (Cabinet Office, 2014). This birth cohort enables us to use longitudinal data to examine how clustering is influenced by early-life factors and examine consequences in middle age. We focused on risk behaviours at age 16 so that we could examine multiple behaviours, some of which manifest only with any prevalence in mid-adolescence (such as sexual activity and drug use), and because of the greater availability of indicators of risk behaviours at this wave.

Our analysis is informed by hypotheses derived from the social development model (Hawkins & Weis, 1985). This proposes that behaviours are learnt in social settings, being a function of observation of others, opportunity for enactment and positive reinforcement by others. This model would suggest that the most important early-life causes of premature involvement in adult behaviours and involvement in anti-social behaviours might relate to family structure and functioning, whereas SES might be more important as an influence on physical inactivity, which is supported by existing evidence (Stalsberg & Pedersen, 2010). We

hypothesise that we will identify several classes of adolescent risk behaviour: low involvement; premature involvement in adult behaviours (smoking, alcohol use, sexual debut); involvement in anti-social behaviours (drug use, fights, delinquency, unprotected sex); and involvement in other risks, informed by previous studies (Lippke et al., 2012). We also hypothesise that adolescent involvement in adult behaviours and anti-social behaviours will associate most strongly with family structure and functioning, as well as early-life indicators of risk such as externalising behaviours, whereas involvement in other risk will be associated most strongly with SES. We hypothesise that adolescent premature involvement in adult behaviours and anti-social behaviours will be more associated with more risk and morbidity in middle age than will be adolescent involvement in physical inactivity alone. We hypothesise that associations with worse outcomes in middle age will be stronger for females and those of lower SES, reflecting the interaction of health-related risk factors and broader biological and socio-economic vulnerabilities.

2. Methods

2.1. Study population and instruments for relevant data collection

The 1970 British Cohort Study (BCS70) follows those born in one week in 1970 in England, Wales and Scotland. Including the birth survey, there have been nine sweeps up to September 2018 (Elliott & Shepherd, 2006). We used data from the birth survey, the second (age 5), the third (age 10), the fourth (age16) and the ninth (age 42) sweep. At the birth survey, data were collected from midwife-reports and clinical records. Afterwards, information was collected through a home interview (age 5), a maternal self-completion questionnaire (age 5), a test booklet (age 5), a parental interview and self-completion questionnaire (age 10 and 16), a medical examination (age 10 and 16), educational assessments (age 10 and 16), a cohort-member self-completion questionnaire (age 10 and 16), and a teacher and head-teacher questionnaire (age 10 and 16). At the ninth sweep, interviews comprising computer-assisted

personal interviews and computer-assisted self-interviews were undertaken. We excluded those who did not answer any relevant questions regarding risk behaviours at age 16. Our study sample was 8754 participants (53.3% women) (see Appendix 1, online supplementary file). The original questionnaires are presented in Appendix 2, and the summary tables of variables are found in Appendices 3 to 5.

2.2. Measures

2.2.1. Multiple risk behaviours at age 16

We examined seven domains of behaviours via cohort-member questionnaires: smoking status and number of cigarettes in a week; binge drinking and amount of alcohol; drug use; early sexual intercourse; unprotected sex; physical inactivity; and involvement in fights and delinquency. Smoking status was created as binary (smoker/non-smoker). The 'Number of cigarettes' variable was re-categorised into five categories (non-smoker, 1-5, 5-20, 21-40, 41+/week). Binge drinking and number of days drinking were re-categorised from an original variable (binge drinking; never, once, more than once, number of days drinking; never, 1-2, 3+/week). Regarding drug use, we created one binary variable from six original variables in relation to use of uppers, downers, cannabis, LSD, cocaine and heroin (yes/no). With respect to the sexual debut before 16, we used an original binary variable (yes/no) in the analysis. In relation to unprotected sex, we derived a binary variable (yes/no) from two variables relevant to contraceptive methods; boys' withdraw and trusting luck. As for physical inactivity and delinquency, we employed the original binary (yes/no) variable. Regarding violence, a binary variable (yes/no) was derived from five original variables which were having taken part in fights at school, at football matches, outside pub, on bus/train and elsewhere.

2.2.2. Risk behaviours and health outcomes at age 42

We employed measurements of various behaviours and health outcomes from the questionnaire-based data collected at age of 42. To examine association between multiple risk-

taking in adolescence and in adulthood, “smoking status (never-smoked, ex-smokers, occasional smokers, every day smokers)”, “number of cigarettes a day”, “number of units of alcohol consumed in the last week” and “number of days a week does 30 minutes or more exercise” were derived. With regard to exercise, we categorised the original variable into binary variable; “none”, and “once and more a week”. In relation to the health outcomes, the measures of “self-rated health (excellent, very good, good, fair, poor)”, “BMI” and “Malaise inventory score (Rutter, Tizard, & Whitmore, 1970), a measure of psychological distress (0-3, 4+)” were employed.

2.2.3. Potential predictors of clustering of risk behaviours

As individual characteristics, we included sex reported at age 16. We derived from the births survey: birthweight (grams), mothers’ age at delivery and mothers’ smoking during pregnancy (smoked/not smoked during the entire period of pregnancy). From the third sweep (age 10), we derived: BMI, medical-examination-based health condition, cognitive ability, behaviour difficulties (Rutter Behaviour Scales (Rutter et al., 1970)), locus of control (CARALOC) (Gammage, 1974) and self-esteem (LAWSEQ) (Lawrence, 1981). The medical examination assessed fifteen conditions (e.g., cardiovascular, neurological). From these, we created a categorical variable composed of “none”, “one”, “more than one” abnormal condition. Cognitive ability was measured by eight tests: shortened Edinburgh Reading Test (Godfrey Thompson Unit, 1978); Friendly Maths test; Pictorial-Language Comprehension test; spelling-dictation task; and British Ability Scales (BAS) comprising four tests (Elliot, Murray, & Pearson, 1979). To obtain individual scores across these eight tests, principal component analysis was applied (Schoon, 2010). We identified a principal component from the total score of eight tests, accounting for 58% of variance. The standardised score of this principal component was used in the analysis. Rutter Behaviour Scales assessed behavioural difficulties drawing on 19 items (Rutter et al., 1970), completed by the cohort-members’ parents. We

identified the first two principal components, accounting for approximately 34% of variance from the total scores. Although the explained variance is relatively small, the eigenvalue of the first component is 4.71, and that of the second component is 1.67. Therefore, we employed these components in the analysis as indicators of externalising (first component) and internalising (second component) behaviour problems, as per previous research (McCulloch, Wiggins, Joshi, & Sachdev, 2000). Externalising behaviour problems include difficulties interacting with other children and at home, difficulties in concentration, having a strong temper and being argumentative. Internalising behaviour problems include being withdrawn, demanding attention, being too dependent or clingy and feeling worthless or inferior (McCulloch et al., 2000). Regarding the locus of control, we used the total score of fifteen items from CARALOC, ranging 0-15 developed by Gammage (Gammage, 1974), completed by cohort-members. Self-esteem was measured using 12 items of Lawrence's self-esteem questionnaire completed by cohort-members (LAWSEQ), ranging 0- 12 (Lawrence, 1981). Previous literatures (Hair et al., 2009; Hale & Viner, 2016) suggest also considering age and ethnicity. Nevertheless we did not include these variables, because the cohort members of BCS70 were born in a single week of 1970 and less than 4% of cohort-members reported non-White ethnicity.

Regarding family characteristics, we employed parental socioeconomic status from the births survey, and family structure, parental smoking and mothers' malaise inventory score from the third sweep, and parental awareness and family rules from the fourth sweep. Parental socioeconomic status was categorised as "professional", "managerial and technical", "skilled-non manual", "skilled-manual", "partly skilled" and "unskilled". Family structure was categorised as: "living with both biological parents", "living with one biological and one non-biological parent", "single parent", and "others". Parental smoking was categorised as "non-smokers", "one parent being a smoker", and "both smokers". Mothers' malaise inventory

score is a continuous variable derived from 24 ranging 0- 100. For parental awareness and family rules, we obtained total scores from twenty-three binary yes/no items.

2.3. Analyses

2.3.1. Latent class analysis (analysis 1)

Latent class analysis aims to describe exhaustive and probabilistic unmeasured classes/groups (Henry & Muthén, 2010) using categorical and/or continuous observed variables (Collins & Lanza, 2009; Vermunt, 2003). We used a combination of model indices; scaled relative entropy (Celeux & Soromenho, 1996), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and sample-size-adjusted BIC to determine the optimal number of latent classes to address our first research question. We aimed to inspect conditional probabilities of the risk behaviours that form the latent classes to distinguish between classes and describe their makeup, and to describe separation between classes and overall model fit as a way of understanding the degree to which latent classes represent meaningful clustering. We conducted analyses stratified by sex, which allowed us to have different conditional probabilities in latent classes across sex. We regressed the conditional probabilities of latent class against sex to examine if there was a difference in probabilities by sex. To describe latent classes in adolescence, we employed the measurements of risk behaviours at age 16 (fourth sweep) on seven domains: smoking status and number of cigarettes in a week; binge drinking and amount of alcohol; drug use; early sexual intercourse; unprotected sex; physical inactivity; and involvement in fights and delinquency.

2.3.2. Regression analysis (analysis 2)

We conducted multivariable logistic regression analyses to investigate our second question about what characteristics, derived from the fourth sweep (age 16) and earlier sweeps, predict latent class membership in adolescent risk behaviours. For our third question, we tested the relationship between latent-class membership and adverse outcomes at age 42

in a separate model with each outcome. We applied multivariable regression modelling depending on the nature of each outcomes: multinomial regression for smoking status; Poisson regression for number of cigarettes; quintile regression for alcohol units; ordinal logistic regression for self-rated health; linear regression for BMI; and logistic regression for physical activity and the malaise inventory score. In these analyses, we applied two models: an unadjusted model and an adjusted model with birth weight, mother's age at delivery, maternal smoking during pregnancy, BMI, abnormal health conditions, cognitive ability, locus of control, self-esteem, Rutter Behaviour Scales score, family structure, parental SES at birth, parental smoking, parental awareness and family rules, and maternal malaise score. To test whether parental SES at birth modified the association between multiple adolescent risk behaviours and outcomes later in life, we used an interaction term in each regression model, and applied Wald tests to determine significance.

2.3.3. Missing data

Of the sample of 17 195 at the birth survey, 11 615 participated in the fourth sweep at age 16 (67.5%), where 8754 responded to at least one question from which the latent classes were derived. There was a substantial number of missing values and incomplete data in confounders derived from earlier sweeps. Regarding outcomes, 69.4% of the 8754 cohort members responded to the ninth sweep, from which outcomes data were derived. Consequently, the number of complete case analysis was 1305 (Appendix 1, online supplementary file).

We applied multiple imputation by chained equations for those who responded to at least one of the questions from which latent classes were derived. This is a fully conditional specification in imputation which does not rely on the assumption of joint multivariate normal distribution. Regression equations were specified for each variable, and imputations performed through estimation of each conditional distribution using observed data (Lee &

Carlin, 2010). We specified each regression model to generate imputations. Regarding the non-normally distributed continuous variables, predictive mean matching was applied. We created 50 imputed datasets since the maximum proportion of incomplete cases in our data was approximately 50% (White, Royston, & Wood, 2011).

All measurements in the analyses, including indicator variables of latent class, potential confounders and outcomes, were included in the imputation model. We also added variables collected at age 42 for longstanding illness, employment grades and marital status as auxiliary variables to increase the probability of meeting the missing-at-random assumption. Previous analyses of cohort attrition suggest that these factors may predict or be associated with missingness (Mein et al., 2012). Moreover, these characteristics are more likely to be associated with health outcomes (Ferrie et al., 2009; Jousilahti, Salomaa, Kuulasmaa, Niemela, & Vartiainen, 2005). The multiply imputed data have stronger statistical power and are more robust due to the use of auxiliary variables in comparison with the complete-case analysis. Therefore, we present and discuss the results based on the multiply imputed data in the text, although we still report the results of the complete-case analysis in online supplementary file. We used Mplus ver.7 for the analysis 1, and Stata SE ver.15.1 for the analysis 2.

2.3.4. Sensitivity analysis

We employed an E-value approach to probe no unmeasured/residual confounding assumption (Ding & VanderWeele, 2016; VanderWeele & Ding, 2017). This approach allows researchers, with minimal assumptions, to calculate the strength of confounding that is needed to account for findings. A large E-value implies that considerable unmeasured confounding may be needed. The formula for E-value can be found elsewhere (VanderWeele & Ding, 2017).

3. Results

3.1. Sex and parental socioeconomic status (SES) of study sample

Our study sample consisted of 8754 individuals (53.1% women). Parental SES at birth was: 5.9% professional, 15.0% managerial/technical, 14.6% skilled-non manual, 44.1% skilled-manual, 15.2% partly skilled and 5.1% unskilled employment.

3.2. Latent class of risk behaviours at age 16

Model fit declined as the number of latent classes increased (Table 1). Therefore, we selected two latent classes, the entropy of which was 0.878: a 'risky-behaviour' class (men: n=820, 20.0%; women: n=1101, 23.6%) and a 'less-risky-behaviour' class (men: n=3270, 80.0%; women: n=3563, 76.4%). However, there was a statistically significant difference in the estimated probabilities across sex (coefficient=-0.182, p-value=0.004). Figure 1 presents the conditional probabilities of the sex-specific two-class solution. Thus, we conducted analysis 2, stratified by sex. Among men, those in the risky-behaviour class were more likely to report smoking (98.2%), binge-drinking (72.0%), sexual debut before 16 (76.6%), fights (44.3%), and delinquency (30.2%). Women in the risky-behaviour class were less likely to engage in binge-drinking (59.6%), fights (16.1%), and delinquency (12.6%). Those in the less-risky-behaviour class generally reported less involvement in these behaviours, although the difference for physical inactivity was minimal and for drug use, unprotected sex, fights and delinquency were small.

3.3. Predictors of class membership

Results of descriptive analyses are presented in Appendix 6, online supplementary file. Table 2 shows the associations of each covariate with the latent class. Taking multiple comparisons into account, earlier involvement in externalising behaviours was associated with higher probability of being in the risky-behaviour class (men: odds Ratio (OR) 1.12, 95% confidence Interval (CI) 1.07, 1.18; women: OR 1.13, 95% CI 1.07, 1.19). Not living with both biological parents was associated with higher probability of being in the risky-behaviour class, whereas parental monitoring and family rules were protective against involvement in risky

behaviours (men: OR 0.92, 95% CI 0.90, 0.94; women: OR 0.89, 95% CI 0.87, 0.91). Parental smoking was associated with cohort-members' engagement in risky behaviours, and the magnitude of association was larger when both parents were smokers (men: OR 1.44 95%CI 1.11, 1.89; women: OR 1.76, 95% CI 1.38, 2.24) than when only one smoked (men: OR 1.37, 95% CI 1.11, 1.68; women: OR 1.31, 95% CI 1.08, 1.58). Parental SES among men was protective factor, but the strength of association was weak.

3.4. Latent class as a predictor of risk behaviours and health outcomes at age 42

In our sample, 48.6% of all participants were non-smokers, 27.0% ex-smokers, 5.5% occasional smokers, and 18.9% everyday smokers. For alcohol consumption, 39.7% of participants drank more than 14 units during the relevant week, and 28.0% reported that they did not do 30 minutes or more exercise in a typical week. Regarding health outcomes, 22.3% reported "excellent" health, 36.5% "very good", 26.7% "good", 10.4% "fair" and 4.1% reported "poor". Mean BMI was 26.9 (SD 5.3). Regarding psychological distress, 18.5% of participants reported depressive symptoms.

Among men, those in the risky-behaviour class at age 16 were 10.1 times more likely to report being an ex-smoker (relative risk ratio (RRR) 10.13, 95% CI 7.40, 13.88), and 16.7 times more likely to report being an everyday smoker at age 42 (RRR 16.73, 95% CI 12.08, 23.18) compared to non-smokers at age 42 who were in the risky behaviour class in a fully adjusted model (Table 3.1), but there was no difference in the number of cigarettes among smokers. Women in the risky-behaviour class at the age of 16 had 23.0 times higher risk of being everyday smokers (RRR 22.98, 95% CI 16.85, 31.36), and were 13.6 times more likely to be ex-smokers (RRR 13.63, 95% CI 10.42, 17.82) than non-smokers at age 42 who were in the risky-behaviour class (Table 3.2). With regard to the number of cigarettes, female smokers at age 42 who were in the risky-behaviour class at age 16 were likely to smoke more cigarettes per day than smokers who were in the less-risky-behaviour class (Incidence Rate Ratio (IRR)

1.11, 95% CI 1.06, 1.17). In relation to alcohol consumption, men in the risky-behaviour class drank more than those in the less-risky-behaviour class (50th percentile: unstandardised coefficient (b) 4.31, 95% CI 2.01, 6.62; 75th percentile: b 5.91, 95% CI 2.55, 9.26). Likewise, among women, those in the risky-behaviour class at age 16 reported higher units of alcohol consumption compared to those in the less-risky-behaviour class (50th percentile: b 1.68, 95% CI 0.67, 2.70; 75th percentile: b 3.30, 95% CI 1.38, 5.23). With regard to self-rated health, those in the risky-behaviour class at age 16 were more likely to report worse health at age 42 than those in the less-risky-behaviour class (men: OR 1.44, 95 CI 1.22, 1.70; women: OR 1.34, 95% CI 1.15, 1.55). Women (but not men) in the risky-behaviour class at age 16 were more likely to report depressive symptoms in middle age (OR 1.40, 95% CI 1.13, 1.72). There were no or weak associations between being in the risk-behaviour class at age 16 and low physical activity or high BMI among both men and women.

We also examined whether SES at birth moderated the association between being in the risky-behaviour class at age 16 and outcomes at age 42. Parental SES at birth moderated the association of class membership with the reported number of cigarettes smoked at age 42 in the fully adjusted model among women (Table 4), such that the magnitude of association was larger among those of lower SES. With other outcomes and among men, there was no evidence of moderation by parental SES.

Our sensitivity analysis suggested that strong unmeasured confounding would be needed to explain away our findings. For example, for the association between being involved in clustered risk behaviours at 16 and poor self-rated health in midlife among women, a confounder with an OR of 2.01 (95% CI 1.57, 2.47) would be needed to account for our findings. Similar results were obtained with all other outcomes (details available from the corresponding author). The results of the complete-case analysis were presented in Appendix 7, online supplementary file. Due to the large number of missing values, the magnitude of

association differed from the imputed data analysis. The direction of association, however, was similar to the results of the imputed data analysis.

4. Discussion

4.1. Adolescent risk behaviours and lifelong health

The findings of the present study advance our understanding of development process for risk behaviours from childhood to adolescence, and the health consequences later in life. Building on previous studies, we found a high prevalence of clustered risk behaviours (Daw et al., 2017; Sanchez-Oliva et al., 2018; Yu et al., 2017). As suggested in some previous studies, involvement was slightly higher among females as well as among those raised in families with low parental monitoring (Daw et al., 2017; DuRant et al., 1999; Pulvers et al., 2018). Our research also suggests earlier involvement in externalising behaviours, and not living with both biological parents as important predictors of engagement in risk behaviours in adolescence. Furthermore, our study for the first time shows the consequences of clustered risk behaviours in adolescence for increased risk and worse health in middle age with some increased risks being particularly strong among women and those of lower SES, extending previous studies which have examined consequences only in early adulthood (Champion et al., 2018; El Ansari et al., 2018). These previous studies have documented that engagement in risk behaviours in early life is more likely to be associated with worse health outcomes and academic performance (Champion et al., 2018; El Ansari et al., 2018). By synthesising these results with our findings in middle age, we could hypothesise that early life engagement in risk behaviours is a determinant of adverse health trajectories across life, although there are various challenges to examine the causal pathways in observational studies (Ben-Shlomo, 2002; Kuh, Ben-Shlomo, Lynch, Hallqvist, & Power, 2003). Moreover, given that both the extent and consequences of risk clustering are influenced by socio-demographic and family factors, risk

clustering is likely to be an important driver of health inequalities. However, most prevention interventions address discrete outcomes, and so may fail to disrupt the mechanisms by which risks cluster and consolidate each other. There have been increasing calls for interventions aiming to address multiple risks (Catalano, Hawkins, Berglund, Pollard, & Arthur, 2002; Cooper, Lhussier, Shucksmith, & Carr, 2017; Knight et al., 2017) emphasising both impact and feasibility (Bonell et al., 2014; D.R. Hale, Fitzgerald-Yau, & Viner, 2014). For policy development, further research is required to examine the pattern, causes and consequences of risk clustering in subsequent cohorts and in other countries.

4.2. Limitations

The cohort was subject to attrition, but this was accounted for by our use of multiple imputation. Our cohort was born in 1970, so our results only shed light on the influences on and consequences of clustered risk behaviours of those now in middle age. We did not examine risk behaviours in the domains of diet, self-harm or accidental injury. Our analysis of health in middle age was limited to self-rated physical and mental health and BMI, and did not examine clinical morbidity or mortality. BCS70 is an observational longitudinal study, and bias due to unmeasured confounding/omitted variables cannot be ruled out. Sensitivity analysis, however, indicated that strong confounding, in most cases stronger than that observed in our data, would be needed to explain our findings. Also, due to the difference of types/categories of data collected across sweeps in BCS70, we were unable to conduct analysis, such as latent-transition analysis, across the seven domains of risk behaviours to examine the association of risk behaviours in early- and late-adolescence, with the health outcomes in middle-age. As it has been documented that risk behaviours in earlier age are more likely to be harmful on health, it is interesting to investigate the pattern of transition of risk behaviours and their association with health outcomes in middle-age in future research (Lanza & Cooper, 2016).

References

Alamian, A., & Paradis, G. (2012). Individual and social determinants of multiple chronic disease behavioral risk factors among youth. *BMC Public Health*, 224.

Bellis, M. A., Hughes, K., Calafat, A., Juan, M., Ramon, A., Rodriguez, J. A., . . . Phillips-Howard, P. (2008). Sexual uses of alcohol and drugs and the associated risks: a cross-sectional study of young people in nine European studies. *BMC Public Health, 8*, 155.

Ben-Shlomo, Y. (2002). A life course approach to chronic disease epidemiology: conceptual models, empirical challenges and interdisciplinary perspectives. *International Journal of Epidemiology, 31*(2), 285-293. doi:10.1093/ije/31.2.285

Cabinet Office. (2014). *Risk Behaviours and Negative Outcomes.* London Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/452169/data_pack_risk_behaviours_and_negative_outcomes.pdf

Celeux, G., & Soromenho, G. (1996). An entropy criterion for assessing the number of clusters in a mixture model. *Journal of Classification, 13*(2), 195-212. doi:10.1007/bf01246098

Champion, K. E., Mather, M., Spring, B., Kay-Lambkin, F., Teesson, M., & Newton, N. C. (2018). Clustering of Multiple Risk Behaviors Among a Sample of 18-Year-Old Australians and Associations With Mental Health Outcomes: A Latent Class Analysis. *Front Public Health, 6*, 135. doi:10.3389/fpubh.2018.00135

Collins, L. M., & Lanza, S. T. (2009). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences* (Vol. 718): John Wiley & Sons.

Connell, C. M., Gilreath, T. D., & Hansen, N. B. (2009). A Multiprocess Latent Class Analysis of the Co-Occurrence of Substance Use and Sexual Risk Behavior Among Adolescents. *70*(6), 943-951. doi:10.15288/jsad.2009.70.943

Daw, J., Margolis, R., & Wright, L. (2017). Emerging Adulthood, Emergent Health Lifestyles: Sociodemographic Determinants of Trajectories of Smoking, Binge Drinking, Obesity, and Sedentary Behavior. *J Health Soc Behav, 58*(2), 181-197. doi:10.1177/0022146517702421

Ding, P., & VanderWeele, T. J. (2016). Sensitivity analysis without assumptions. *Epidemiology, 27*(3), 368.

DuRant, R. H., Smith, J. A., Kreiter, S. R., & Krowchuk, D. P. (1999). The relationship between early age of onset of initial substance use and engaging in multiple health risk behaviors among young adolescents. *Arch Pediatr Adolesc Med, 153*(3), 286-291.

El Ansari, W., Ssewanyana, D., & Stock, C. (2018). Behavioral Health Risk Profiles of Undergraduate University Students in England, Wales, and Northern Ireland: A Cluster Analysis. *Front Public Health, 6*, 120. doi:10.3389/fpubh.2018.00120

Elliot, C., Murray, D., & Pearson, L. (1979). British ability scales, manual 4: Tables of abilities and norms. *Windsor: NFER*.

Elliott, J., & Shepherd, P. (2006). Cohort profile: 1970 British Birth Cohort (BCS70). *Int J Epidemiol, 35*(4), 836-843. doi:10.1093/ije/dyl174

Fergusson, D. M., Horwood, L. J., & Lynskey, M. T. J. J. o. A. C. P. (1994). The comorbidities of adolescent problem behaviors: A latent class model. *22*(3), 339-354. doi:10.1007/bf02168078

Ferrie, J. E., Kivimaki, M., Singh-Manoux, A., Shortt, A., Martikainen, P., Head, J., . . . Shipley, M. J. (2009). Non-response to baseline, non-response to follow-up and mortality in the Whitehall II cohort. *International Journal of Epidemiology, 38*(3), 831-837. doi:10.1093/ije/dyp153

Gammage, P. (1974). *Socialisation, schooling and locus of control.* University of Bristol,

Godfrey Thompson Unit. (1978). University of Edinburgh. *Edinburgh Reading Test", Sevenoaks: Hodder and Stoughton*.

Hair, E. C., Park, M. J., Ling, T. J., & Moore, K. A. (2009). Risky behaviors in late adolescence: co-occurrence, predictors, and consequences. *J Adolesc Health, 45*(3), 253-261. doi:10.1016/j.jadohealth.2009.02.009

Hale, D. R., & Viner, R. M. (2016). The correlates and course of multiple health risk behaviour in adolescence. *BMC Public Health, 16*, 458. doi:10.1186/s12889-016-3120-z

Hart, C. L., Morrison, D. S., Batty, G. D., Mitchell, R. J., & Davey Smith, G. (2010). Effect of body mass index and alcohol consumption on liver disease: analysis of data from two prospective cohort studies. *Bmj, 340*, c1240. doi:10.1136/bmj.c1240

Hawkins, J. D., & Weis, J. G. (1985). The social development model: An integrated approach to delinquency prevention. *J Prim Prev, 6*(2), 73-97. doi:10.1007/bf01325432

Henry, K. L., & Muthén, B. (2010). Multilevel latent class analysis: An application of adolescent smoking typologies with individual and contextual predictors. *Structural Equation Modeling, 17*(2), 193-215.

Jackson, C., Sweeting, H., & Haw, S. (2012). Clustering of substance use and sexual risk behaviour in adolescence: analysis of two cohort studies. *BMJ Open, 2*, e000661. doi:10.1136/bmjopen-2011-000661

Jackson, C. A., Henderson, M., Frank, J. W., & Haw, S. J. (2012). An overview of prevention of multiple risk behaviour in adolescence and young adulthood. *Journal of Public Health, 34*(S1), i34-i40.

Jousilahti, P., Salomaa, V., Kuulasmaa, K., Niemela, M., & Vartiainen, E. (2005). Total and cause specific mortality among participants and non-participants of population based health surveys: a comprehensive follow up of 54 372 Finnish men and women. *J Epidemiol Community Health, 59*(4), 310-315. doi:10.1136/jech.2004.024349

Kuh, D., Ben-Shlomo, Y., Lynch, J., Hallqvist, J., & Power, C. (2003). Life course epidemiology. *J Epidemiol Community Health, 57*(10), 778-783. doi:10.1136/jech.57.10.778

Lanza, S. T., & Cooper, B. R. (2016). Latent Class Analysis for Developmental Research. *10*(1), 59-64. doi:10.1111/cdep.12163

Lawrence, D. (1981). The development of a self‐esteem questionnaire. *British Journal of Educational Psychology, 51*(2), 245-251.

Lee, K. J., & Carlin, J. B. (2010). Multiple imputation for missing data: fully conditional specification versus multivariate normal imputation. *Am J Epidemiol, 171*(5), 624-632. doi:10.1093/aje/kwp425

Lippke, S., Nigg, C. R., & Maddock, J. E. (2012). Health-promoting and health-risk behaviors: theory-driven analyses of multiple health behavior change in three international samples. *Int J Behav Med, 19*(1), 1-13. doi:10.1007/s12529-010-9135-4

McAloney, K. (2015). Clustering of Sex and Substance Use Behaviors in Adolescence. *Subst Use Misuse, 50*(11), 1406-1411. doi:10.3109/10826084.2015.1014059

McCulloch, A., Wiggins, R. D., Joshi, H. E., & Sachdev, D. (2000). Internalising and externalising children's behaviour problems in Britain and the US: relationships to family resources1. *Children & Society, 14*(5), 368-383. doi:doi:10.1111/j.1099-0860.2000.tb00191.x

Meader, N., King, K., Moe-Byrne, T., Wright, K., Graham, H., Petticrew, M., . . . Sowden, A. J. (2016). A systematic review on the clustering and co-occurrence of multiple risk behaviours. *BMC Public Health, 16*, 657. doi:10.1186/s12889-016-3373-6

Mein, G., Johal, S., Grant, R. L., Seale, C., Ashcroft, R., & Tinker, A. (2012). Predictors of two forms of attrition in a longitudinal health study involving ageing participants: an analysis based on the Whitehall II study. *BMC Med Res Methodol, 12*, 164. doi:10.1186/1471-2288-12-164

Meng, L., Maskarinec, G., Lee, J., & Kolonel, L. N. (1999). Lifestyle factors and chronic diseases: application of a composite risk index. *Preventive medicine, 29*(4), 296-304.

Patton, G., Sawyer, S. M., Santelli, J. S., Ross, D. A., Afifi, R., Allen, N. B., . . . Viner, R. (2016). Realising the Future of Global Health: A Lancet Commission on Adolescent Health and Wellbeing. *The Lancet, 387*(10036), 2423-2478.

Prabhu, A., Obi, K. O., & Rubenstein, J. H. (2014). The synergistic effects of alcohol and tobacco consumption on the risk of esophageal squamous cell carcinoma: a meta-analysis. *American Journal of Gastroenterology, 109*, 822-827.

Pulvers, K., Ridenour, C., Woodcock, A., Savin, M. J., Holguin, G., Hamill, S., & Romero, D. R. (2018). Marijuana use among adolescent multiple tobacco product users and unique risks of dual tobacco and marijuana use. *Drug Alcohol Depend, 189*, 80-89. doi:10.1016/j.drugalcdep.2018.04.031

Qin, L., Knol, M. J., Corpeleijn, E., & Stolk, R. P. (2010). Does physical activity modify the risk of obesity for type 2 diabetes: a review of epidemiological data. *Eur J Epidemiol, 25*(1), 5-12. doi:10.1007/s10654-009-9395-y

Rutter, M., Tizard, J., & Whitmore, K. (1970). *Education, health and behaviour*: Longman Publishing Group.

Sanchez-Oliva, D., Grao-Cruces, A., Carbonell-Baeza, A., Cabanas-Sanchez, V., Veiga, O. L., & Castro-Pinero, J. (2018). Lifestyle Clusters in School-Aged Youth and Longitudinal Associations with Fatness: The UP&DOWN Study. *J Pediatr*. doi:10.1016/j.jpeds.2018.07.092

Schoon, I. (2010). Childhood cognitive ability and adult academic attainment: Evidence from three British cohort studies. *Longitudinal and Life Course Studies, 1*(3), 241-158.

Stalsberg, R., & Pedersen, A. V. (2010). Effects of socioeconomic status on the physical activity in adolescents: a systematic review of the evidence. *Scandinavian journal of medicine & science in sports, 20*(3), 368-383.

Tomczyk, S., Isensee, B., & Hanewinkel, R. (2016). Latent classes of polysubstance use among adolescents-a systematic review. *Drug Alcohol Depend, 160*, 12-29. doi:10.1016/j.drugalcdep.2015.11.035

VanderWeele, T. J., & Ding, P. (2017). Sensitivity Analysis in Observational Research: Introducing the E-Value. *Ann Intern Med, 167*(4), 268-274. doi:10.7326/m16-2607

Vermunt, J. K. (2003). Multilevel Latent Class Models. *Sociological Methodology, 33*(1), 213-239. doi:10.1111/j.0081-1750.2003.t01-1-00131.x

Wagner, F. A., & Anthony, J. C. (2002;). Into the world of illegal drug use: exposure opportunity and other mechanisms linking the use of alcohol, tobacco, marijuana, and cocaine. *Am J Epidemiol, 155*(10), 918-925.

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: issues and guidance for practice. *Stat Med, 30*(4), 377-399.

Yu, J., Putnick, D. L., Hendricks, C., & Bornstein, M. H. (2017). Health-Risk Behavior Profiles and Reciprocal Relations With Depressive Symptoms From Adolescence to Young Adulthood. *J Adolesc Health, 61*(6), 773-778. doi:10.1016/j.jadohealth.2017.07.002