Preventive Medicine 88 (2016) 95–107

Contents lists available at ScienceDirect

Preventive Medicine

journal homepage: www.elsevier.com/locate/ypmed

The clustering of health-related behaviours in a British population sample: Testing for cohort differences

Claire Mawditt [a,*], Amanda Sacker [a], Annie Britton [b], Yvonne Kelly [a], Noriko Cable [a]

[a] International Centre for Lifecourse Studies in Society and Health, Research Department of Epidemiology and Public Health, University College London, 1-19 Torrington Place, London, WC1E 6BT, United Kingdom
[b] Research Department of Epidemiology and Public Health, University College London, 1-19 Torrington Place, London, WC1E 6BT, United Kingdom

A R T I C L E I N F O

Article history:
Received 10 November 2015
Received in revised form 10 February 2016
Accepted 6 March 2016
Available online 4 April 2016

Keywords:
Latent Profile Analysis
Latent Class Analysis
Clustering
Health-related behaviour
Alcohol
Smoking
Diet
Physical activity
Cohort
Gender

A B S T R A C T

Research findings indicate that health-related behaviours (HRBs) do not co-occur within individuals by chance and therefore cluster. This study uses Latent Profile Analysis (LPA), to identify the clustered patterns and prevalence of four HRBs: smoking, alcohol, diet, physical activity. We used data, collected from participants in their early 30s, from two British cohorts born in 1958 and 1970 (N = 21,019). Multi-group LPA models were run separately by gender testing for cohort differences in HRB cluster patterns. For both genders three clusters emerged: 'Risky' (1–9%), 'Moderate Smokers' (20–30%) and 'Mainstream' (68–77%). HRBs amongst members of the 'Mainstream' cluster were more beneficial than HRBs amongst members of the other two clusters, characterised as not smoking, frequent fruit and vegetable consumption, less frequent consumption of chips and fried food and being more physically active. Nevertheless, frequent consumption of sweet foods was common in the 'Mainstream' cluster. There was a large shift in membership to the 'Mainstream' cluster for men and women born in 1970. Amongst women members of the 'Mainstream' cluster, a higher proportion of those born in 1970 appeared to have drunk alcohol above the contemporaneous UK recommended limits but consumed sweet foods less frequently, than those born in 1958. In summary our findings provide additional evidence of HRB clustering, identifying largely consistent HRBs cluster patterns across cohort and gender groups, with some differences in prevalence. This evidence of HRB clustering across time and by gender provides a person-centred understanding that can inform interventions to improve HRBs.

1. Introduction

Modifiable negative health related behaviours (HRBs) such as smoking, heavy alcohol consumption, physical inactivity and an unhealthy diet, characterised as high in sugar and fat, low in fruit and vegetables are leading causes of non-communicable disease globally (WHO, 2014), and strongly associated with early mortality (Loef and Walach, 2012; Kvaavik et al., 2010).

Evidence from studies using large population based samples, show that individuals commonly have two or more negative HRBs (Schuit et al., 2002; Berrigan et al., 2003; Poortinga, 2007; Silva et al., 2013; Linardakis et al., 2013; Buck and Frosini, 2012; Laaksonen et al., 2001). Of these it appears that there are population subgroups who share distinct behavioural patterns of smoking, alcohol consumption, diet and physical activity (Schuit et al., 2002; Berrigan et al., 2003; Poortinga, 2007; Silva et al., 2013; Laaksonen et al., 2001; van Nieuwenhuijzen et al., 2009; De Vries et al., 2008; Conry et al., 2011; Schneider et al., 2009; Vermeulen-Smit et al., 2015; Leventhal et al., 2014; Heroux et al., 2012; Tobias et al., 2007; Verger et al., 2009), unlikely to be present by chance (Schuit et al., 2002; Berrigan et al., 2003; Poortinga, 2007; Silva et al., 2013). Research investigating HRB clustering has found that people who smoke cigarettes are more likely to drink alcohol heavily and less likely to consume fruit and vegetables and be physically active (Berrigan et al., 2003; Poortinga, 2007; De Vries et al., 2008).

Previous work comparing the HRBs of people born in 1958 and 1970 (Schoon and Parsons, 2003; Elliott et al., 2007) found that some HRBs had improved amongst those born in 1970 e.g. eating chips less frequently and fewer women smokers (Schoon and Parsons, 2003). However, some HRBs were found to be worse e.g. eating fruit less frequently and increased alcohol consumption amongst women born in 1970 (Schoon and Parsons, 2003; Elliott et al., 2007).

Data reduction techniques are useful in identifying clustering patterns (Hofstetter et al., 2014; McAloney et al., 2013). Application of these techniques on HRBs of cohort participants, born 12 years apart, can provide empirical evidence as to whether HRBs cluster (McAloney et al., 2013). Here taking a data driven approach we examine cohort differences of HRB clustering in two ways; cluster patterns (i.e.

* Corresponding author.
E-mail addresses: claire.mawditt.11@ucl.ac.uk (C. Mawditt), a.sacker@ucl.ac.uk (A. Sacker), a.britton@ucl.ac.uk (A. Britton), y.kelly@ucl.ac.uk (Y. Kelly), n.cable@ucl.ac.uk (N. Cable).

http://dx.doi.org/10.1016/j.ypmed.2016.03.003

combinations of HRBs) and cluster membership (i.e. the proportion belonging to each cluster).

Evidence suggests that HRBs across cohorts differ according to gender (Schoon and Parsons, 2003; Elliott et al., 2007). Therefore, cohort differences are examined separately for men and women. Our particular focus is HRBs in mid-adulthood given evidence that HRBs are sustained during this period (Mulder et al., 1998; Harrington et al., 2014; Telama, 2009; Parsons et al., 2006).

2. Methods

2.1. Sample

This study uses cross-sectional data from two British birth cohort studies: The National Child Development Study (NCDS), targeting 17,514 individuals born in the same week in 1958 (Power and Elliott, 2006), and the British Birth Cohort Study (BCS70) targeting 16,571 individuals born in one week in 1970 (Elliott and Shepherd, 2006), across England, Scotland and Wales. The two cohorts are purposefully similar in their design, allowing for a meaningful comparison (Ekinsmyth et al., 1992). Informed consent was obtained by participants agreeing to be interviewed and completing questionnaires, after receiving information on the study and the choice to opt out (Shepherd, 2012a; Shepherd, 2012b).

We used NCDS data collected in 1991, when participants were age 33. Data were available for 11,407 participants, response rate 73%. Participants with complete information on at least one HRB were analysed, yielding a sample of 11,373 (99.7%), 5586 men and 5787 women. We used BCS70 data collected in 2004, when participants were age 34. Data collected in 2000, when participants were age 30, supplemented information about their diet (unavailable at age 34). A total of 9665 participants were included at age 34, response rate 75%. Participants with information on at least one HRB, yielded a sample of 9646 (99.8%), 4613 men and 5033 women.

2.2. Measures

Four HRBs: smoking, alcohol, diet and physical activity were measured using six variables: cigarette smoking, alcohol unit consumption, fruit and vegetable consumption, chips and fried food consumption, sweets, chocolate, biscuits and cakes consumption, and physical activity frequency. The alcohol measure is based upon UK government guidelines active in 1991 and 2004 for 'safe' weekly consumption (DOH. Sensible drinking: Report of an inter-departmental working group. London: Department of Health, 1995). The measures of smoking, diet and physical activity are pragmatically determined. Appendix A describes the questionnaire items and cohort harmonisation.

Participants were asked if they smoked cigarettes and the average number smoked per day (range 0–80) those who reported not smoking cigarettes were coded as 0. Those reporting to smoke occasionally (BCS70 only, $n = 645$, 6.4%),were also coded as 0.

Alcohol consumption was measured according to average drinking frequency and the number of alcoholic beverages consumed in the previous week. Beverage categories were combined to provide the total number of units consumed (1 unit = 8 g ethanol, range 0–210 units). This total was categorised according to consumption frequency and quantity, reflecting gender specific UK guidelines for 'safe' weekly consumption (DOH. Sensible drinking: Report of an inter-departmental working group. London: Department of Health, 1995). Participants reporting 0 units in the previous week were coded as 'no units' alongside never and infrequent drinkers. Men reporting 1–21 units and women reporting 1–14 units were coded as 'within limits' as were frequent drinkers, reporting 0 units in the previous week. Men reporting >21 units and women reporting >14 units were coded as 'above limits'.

Participants were asked whether they regularly took part in leisure time physical activity, defined as "at least once a month, for most of the year", and the frequency; "every day", "4–5 days per week", "2–3 days per week", "once a week", "2–3 times a month", "less often". Responses with sparse data were combined, creating four categories; '≤3 times a month', 'once a week', '2–3 days a week', '4–7 days a week'.

Diet was indicated by the average frequency of consumption of six variables; 'fruit'; 'vegetables'; 'chips'; 'fried foods'; 'sweets or chocolate'; 'biscuits' (NCDS), and 'biscuits or cakes' (BCS70). In both studies, participants were asked if they consumed these foods "more than once a day", "once a day", "3–6 days a week", "1–2 days a week", "less than 1 day a week" or "never". An additional "occasional" category, present in the BCS70, was combined with "less than 1 day a week".

Based on the findings from Principal Components Analysis, the six diet variables were combined to form three composite variables; 'fruit and vegetables', 'chips and fried food' (hereafter fried food) and 'sweets, chocolate, biscuits or cakes' (hereafter sweet food). Frequency scores (range 0 to 5) were added together, creating a score ranging from 0 (never) to 10 (more than once a day).

2.3. Statistical analyses

Latent Profile Analysis (LPA) (Collins and Lanza, 2010) was conducted using Mplus version 7 (Muthen, 2014), to identify HRB clustering. LPA incorporates continuous variables, but is otherwise identical to Latent Class Analysis (Collins and Lanza, 2010), which is increasingly utilised to identify HRB clusters (De Vries et al., 2008; Vermeulen-Smit et al., 2015; Leventhal et al., 2014; Heroux et al., 2012; Schnuerer et al., 2015; Watts et al., 2015) and has formal statistical procedures to guide the selection of clusters (Wang and Wang, 2012). LPA models assume that observed variables are conditionally independent, and that associations between them are explained by the latent (unobserved) variable (Collins and Lanza, 2010). We relaxed this assumption, allowing diet variables to correlate within each cluster.

Preliminary analysis found smoking to be rare in the largest cluster, this variable also had a long right-tailed distribution. To aid model convergence, the mean and variance of smoking was fixed at zero in the largest cluster and the distribution was condensed by dividing the variable by ten. Behaviour variables were continuous or ordered, rather than binary, to retain more information on individual differences in the data. In all models 4000 different starting values were used to identify the maximum likelihood solution (Collins and Lanza, 2010).

To determine an optimal number of clusters, several LPA models were estimated, adding another cluster (k) to each consecutive model and comparing fit indices to the previous model (k-1). Fit indices included the likelihood ratio chi-squared test; entropy; adjusted Bayesian Information Criterion (aBIC); and the Lo–Mendell Rubin likelihood ratio test (LMR) (Collins and Lanza, 2010; Nylund et al., 2007). Emphasis was placed upon the aBIC which balances model fit and parsimony (McAloney et al., 2013; Collins and Lanza, 2010) and minimum cluster size criterion was established (Wang and Wang, 2012). As recommended (Collins and Lanza, 2010), prior to conducting multi-group LPA, models were run separately for each subgroup (NCDS men, BCS70 men, NCDS women, BCS70 women) to establish whether the same number of clusters emerged. This was followed by multi-group LPA models (Collins and Lanza, 2010), run separately for men and women, stratifying the sample according to cohort. Wald chi-square tests were performed to detect differences in HRB means and response probabilities within and across each cohort, for men and women. Wald chi-square tests were used to detect cohort differences in the proportion of participants in each cluster.

Measurement invariance analysis was conducted to assess cluster equivalence across the cohorts (Finch, 2015), described in Appendix B.

Full Information Maximum Likelihood (FIML), utilising all available information in the data under a missing at random (MAR) assumption (Enders, 2010), was employed to manage missing data. HRB means and response probabilities from complete case models were compared with those from FIML models.

3. Results

3.1. Descriptive analysis

For both men and women, behaviours tended to be healthier (e.g. smoking fewer cigarettes per day, higher frequency of fruit and vegetable consumption, higher frequency of physical activity) in the BCS70 compared to the NCDS (Table 1). An exception was alcohol consumption where a higher proportion of participants in the BCS70 drank above recommended limits compared to the NCDS, particularly amongst women.

3.2. Latent Profile Analysis

The aBIC for LPA models run separately for each cohort within each gender, suggested that four clusters were preferred over three. However, for all groups the smallest cluster in the 4 cluster models fell below the minimum cluster size criterion and patterns in the 4th cluster added little to model interpretability. On this basis a 3 cluster multi-group LPA model was chosen for both genders. Model fit indices and minimum cluster size criterion for these models are shown in Appendix C. Estimates from models using FIML are presented below (Tables 2 and 3). The same cluster labels could be assigned across the multi-group models, aiding interpretability. Cluster 1, labelled 'Risky', had patterns riskier than the others (i.e. heavy smoking). Cluster 2, was labelled 'Moderate Smokers', because smoking behaviour notably distinguished this cluster from the others, although levels of smoking were lower than the 'Risky' cluster. Cluster 3, labelled 'Mainstream', was the largest cluster, representing the most prevalent HRB patterns in the data, described below. For both genders, measurement invariance analysis suggested partial cluster equivalence across the cohorts (see Appendix B).

3.3. Cluster patterns

For both genders, cluster patterns were similar across the cohorts for smoking, fruit and vegetable consumption, fried food consumption and physical activity but patterns diverged for sweet food and alcohol consumption.

Wald chi-square tests found the estimated mean number of cigarettes smoked per day was higher for members in the 'Risky' and 'Moderate Smokers' clusters ('Risky' NCDS men = 41 cigarettes, NCDS women = 21 cigarettes; 'Moderate Smokers' NCDS men = 17 cigarettes, NCDS women = 14 cigarettes), compared to those in the 'Mainstream' cluster ($p < 0.01$), which was fixed at zero in line with our methodological approach.

Members of the 'Risky' cluster had lower frequencies of fruit and vegetable consumption (mean NCDS men = 2.61; BCS70 men = 3.75; NCDS women = 3.39; BCS70 women = 3.67) and higher frequencies of fried food consumption (mean NCDS men = 4.73; BCS70 men = 6.73; NCDS women = 4.02; BCS70 women = 3.37) compared to members of the 'Moderate Smokers' and 'Mainstream' clusters ($p < 0.01$).

The frequency of leisure time physical activity was highest for the members of the 'Mainstream' cluster (≥once per week NCDS men = 72%; BCS70 men = 73%; NCDS women = 73%; BCS70 women = 76%), followed respectively by the members of the 'Moderate Smokers' and 'Risky' clusters ($p < 0.01$).

Sweet food consumption frequency was generally highest in the 'Mainstream' cluster and lowest in the 'Risky' cluster ($p < 0.01$). The exception was BCS70 men whose sweet food consumption frequency was high in the 'Mainstream' cluster (mean = 4.59) but highest in the 'Risky' cluster (mean = 5.23, $p < 0.01$). In women, sweet food consumption frequency in the 'Mainstream' cluster was significantly lower amongst

Table 1
Health-related behaviour characteristics of the analytical sample: total pooled and stratified by cohort and gender. Data: two British birth cohort studies, the National Child Development Study (NCDS) at age 33 (1991), the British Cohort Study (BCS70) at age 30 (2000) and 34 (2004).

Health-related behaviour variables	Total Pooled N = 21,019 (100%)	Men NCDS n = 5586 (100%)	Men BCS70 n = 4613 (100%)	Women NCDS n = 5787 (100%)	Women BCS70 n = 5033 (100%)
	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)
Number of cigarettes smoked per day[a]	16.4 (8.5)	18.5 (9.5)	16.0 (7.9)	16.2 (8.2)	13.7 (6.7)
Frequency of fruit and vegetable consumption[b]	5.2 (2.1)	4.5 (1.9)	4.9 (2.1)	5.6 (1.9)	5.8 (2.2)
Frequency of fried food consumption[b]	3.02 (1.3)	3.7 (1.5)	3.0 (1.2)	2.9 (1.3)	2.5 (1.0)
Frequency of sweet food consumption[b]	4.7 (2.2)	4.7 (2.1)	4.6 (2.3)	4.8 (2.2)	4.6 (2.2)
	n (%)	n (%)	n (%)	n (%)	n (%)
Diet Missing	696 (3.31%)	9 (0.2%)	475 (10.3%)	12 (0.2%)	392 (7.8%)
Proportion smoking cigarettes daily					
0	15,022 (71.5%)	3797 (68.0%)	3404 (73.8%)	3964 (68.5%)	3857 (76.6%)
1–10	1934 (9.2%)	458 (8.2%)	385 (8.4%)	573 (9.9%)	518 (10.3%)
11–20	3159 (15.0%)	912 (16.33%)	680 (14.7%)	984 (17.0%)	583 (11.6%)
21+	842 (4.0%)	393 (7.0%)	135 (2.9%)	249 (4.3%)	65 (1.3%)
Missing	62 (0.3%)	26 (0.5%)	9 (0.2%)	17 (0.3%)	10 (0.2%)
Frequency of leisure time physical activity					
≤3 times a month	6300 (30.0%)	1773 (31.7%)	1391 (30.2%)	1775 (30.7%)	1361 (27.0%)
Once a week	4102 (19.5%)	1166 (20.9%)	825 (17.9%)	1314 (22.7%)	797 (15.8%)
2–3 days a week	4932 (23.5%)	1292 (23.1%)	1237 (26.8%)	1110 (19.2%)	1293 (25.7%)
4–7 days a week	5611 (26.7%)	1330 (23.8%)	1156 (25.1%)	1551 (26.8%)	1574 (31.3%)
Missing	74 (0.4%)	25 (0.5%)	4 (0.09%)	37 (0.6%)	8 (0.2%)
Alcohol units consumed in the previous week[c]					
No units	4292 (20.4%)	754 (13.5%)	569 (12.3%)	1670 (28.9%)	1299 (25.8%)
Within limits (≤14 units women, ≤21 units men)	12,484 (59.4%)	3280 (58.7%)	2578 (55.9%)	3640 (62.9%)	2986 (59.3%)
Above limits (≥15 units women, ≥22 units men)	4212 (20.0%)	1549 (27.7%)	1450 (31.4%)	474 (8.2%)	739 (14.7%)
Missing	31 (0.2%)	3 (0.05%)	16 (0.4%)	3 (0.05%)	9 (0.2%)

[a] Range 1–80.
[b] A higher score indicates a higher consumption frequency. Range 0–10. Diet score equivalent (rounded to zero decimal places): 'never' [0] 'occasionally /less than 1 day a week' [1–2] '1–2 days a week' [3–4] '3–6 days a week' [5–6] once a day' [7–8] 'more than once a day' [9–10].
[c] 'No units' category includes never drinkers and non-frequent drinkers who report 0 units in the previous week. Frequent drinkers who report 0 units in the previous week have been placed in category 'within limits'.

Table 2
Estimated means and item response probabilities FIML of 3 cluster multi-group Latent Profile Analysis (LPA) model for men.
Data: Two British birth cohort studies, the National Child Development Study (NCDS) at age 33 (1991), the British Cohort Study (BCS70) at age 30 (2000) and 34 (2004).

	NCDS men n = 5586 (100%)			BCS70 men n = 4613 (100%)		
	Cluster 1 'Risky' n = 82 (1.5%)≠	Cluster 2 'Moderate Smokers' n = 1686 (30.2%)≠	Cluster 3 'Mainstream' n = 3818 (68.3%)≠	Cluster 1 'Risky' n = 79 (1.7%)≠	Cluster 2 'Moderate Smokers' n = 1124 (24.4%)≠	Cluster 3 'Mainstream' n = 3410 (73.9%)≠
	Mean (S.E)	Mean (S.E)	Mean (S.E)	Mean (S.E)	Mean (S.E)	Mean (S.E)
Number of cigarettes smoked per day	40.84 (3.67)*†	17.22 (0.31)*†	0	19.82 (4.46)†	15.57 (0.35)†	0
Frequency of fruit and vegetable consumption	2.61 (0.37)*†	3.95 (0.05)*†	4.64 (0.03)*†	3.75 (0.28)*†	4.29 (0.07)*†	5.10 (0.04)*†
Frequency of fried food consumption	4.73 (0.45)*†	3.99 (0.05)*†	3.36 (0.02)*†	6.74 (0.29)*†	3.02 (0.04)*†	2.86 (0.02)*†
Frequency of sweet food consumption	3.58 (0.45)*	4.18 (0.06)*	4.71 (0.04)*	5.23 (0.53)*	4.34 (0.08)*	4.59 (0.04)*
	Item Response probability (S.E)	Item Response probability (S.E)	Item Response probability (S.E)	Item Response probability (S.E)	Item Response probability (S.E)	Item Response probability (S.E)
Frequency of leisure time physical activity						
≤3 times a month	0.61 (0.07)*	0.39 (0.01)*	0.28 (0.01)*†	0.49 (0.07)*	0.41 (0.02)*	0.26 (0.01)*†
Once a week	0.12 (0.05)	0.21 (0.01)	0.21 (0.01)	0.09 (0.04)	0.18 (0.01)	0.18 (0.01)
2–3 days a week	0.14 (0.04)	0.19 (0.01)	0.25 (0.01)	0.20 (0.05)	0.20 (0.01)	0.29 (0.01)
4–7 days a week	0.13 (0.04)	0.21 (0.01)	0.26 (0.01)	0.23 (0.06)	0.22 (0.01)	0.26 (0.01)
Alcohol units consumed in the previous week						
No units	0.26 (0.08)*	0.14 (0.01)*†	0.13 (0.01)*†	0.26 (0.06)*	0.13 (0.01)*†	0.12 (0.01)*†
Within limits (≤14 units women, ≤21 units men)	0.23 (0.06)	0.50 (0.01)	0.63 (0.01)	0.31 (0.07)	0.44 (0.02)	0.61 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.51 (0.08)	0.36 (0.01)	0.24 (0.01)	0.42 (0.09)	0.43 (0.02)	0.28 (0.01)

Note: * = cluster means and response probabilities are significantly different (p ≤ 0.05) across the three clusters within each cohort. † = cluster means and response probabilities are significantly different (p ≤ 0.01) across the cohorts. ≠ = cluster membership is significantly different (p ≤ 0.01) across the cohorts. Estimated using the Wald chi-square test.

BCS70 members (mean = 4.60) compared to NCDS members (mean = 4.85, $p < 0.01$).

For both genders, alcohol consumption was lowest for members of the 'Mainstream' cluster across cohorts ($p < 0.01$). For men, NCDS members of the 'Risky' cluster had the highest proportion drinking alcohol above recommended limits (51%) compared to the 'Moderate Smokers' cluster (36%), whereas proportions were similar for BCS70 members ('Risky' = 42%; 'Moderate Smokers' = 43%). For women,

Table 3
Estimated means and item response probabilities of FIML 3 cluster multi-group Latent Profile Analysis (LPA) model for women.
Data: Two British birth cohort studies, the National Child Development Study (NCDS) at age 33 (1991), the British Cohort Study (BCS70) at age 30 (2000) and 34 (2004).

	NCDS women total N = 5787 (100%)			BCS70 women total N = 5033 (100%)		
	Cluster 1 'Risky' n = 515 (8.9%)≠	Cluster 2 'Moderate Smokers' n = 1292 (22.3%)≠	Cluster 3 'Mainstream' n = 3980 (68.8%)≠	Cluster 1 'Risky' n = 183 (3.6%)≠	Cluster 2 'Moderate Smokers' n = 984 (19.6%)≠	Cluster 3 'Mainstream' n = 3866 (76.8%)≠
	Mean (S.E)	Mean (S.E)	Mean (S.E)	Mean (S.E)	Mean (S.E)	Mean (S.E)
Number of cigarettes smoked per day	20.96 (1.00)*†	14.07 (0.31)*	0	19.18 (1.88)†	12.30 (0.39)	0
Frequency of fruit and vegetable consumption	3.39 (0.15)*†	5.57 (0.14)*	5.79 (0.03)*†	3.67 (0.16)*†	5.41 (0.20)*	5.97 (0.04)*†
Frequency of fried food consumption	4.02 (0.15)*†	2.69 (0.07)*†	2.55 (0.02)*†	3.37 (0.30)*†	2.32 (0.07)*†	2.36 (0.02)*†
Frequency of sweet food consumption	3.76 (0.24)*	4.40 (0.10)*†	4.85 (0.04)*†	3.68 (0.27)*	4.50 (0.12)*†	4.60 (0.04)*†
	Item Response probability (S.E)	Item Response probability (S.E)	Item Response probability (S.E)	Item Response probability (S.E)	Item Response probability (S.E)	Item Response probability (S.E)
Frequency of leisure time physical activity						
≤3 times a month	0.62 (0.03)*†	0.29 (0.03)*	0.27 (0.01)*†	0.55 (0.06)*†	0.31 (0.03)*	0.25 (0.01)*†
Once a week	0.16 (0.02)	0.21 (0.01)	0.24 (0.01)	0.08 (0.03)	0.15 (0.01)	0.17 (0.01)
2–3 days a week	0.07 (0.02)	0.20 (0.01)	0.21 (0.01)	0.07 (0.03)	0.23 (0.02)	0.28 (0.01)
4–7 days a week	0.15 (0.02)	0.31 (0.01)	0.28 (0.01)	0.30 (0.05)	0.32 (0.02)	0.31 (0.01)
Alcohol units consumed in the previous week						
No units	0.27 (0.03)*	0.30 (0.01)*†	0.29 (0.01)*†	0.40 (0.05)*	0.27 (0.02)*†	0.24 (0.01)*†
Within limits (≤14 units women, ≤21 units men)	0.54 (0.03)	0.61 (0.01)	0.65 (0.01)	0.28 (0.08)	0.54 (0.02)	0.63 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.19 (0.03)	0.09 (0.01)	0.07 (0.01)	0.32 (0.08)	0.18 (0.02)	0.13 (0.01)

Note: * = cluster means and response probabilities are significantly different (p ≤ 0.05) across the three clusters within each cohort. † = cluster means and response probabilities are significantly different (p ≤ 0.01) across the cohorts. ≠ = cluster membership is significantly different (p ≤ 0.01) across the cohorts. Estimated using the Wald chi-square test.

the proportion of BCS70 members drinking above recommended limits was almost double that of NCDS members, across all three clusters ('Risky' NCDS = 19%, BCS70 = 32%; 'Moderate Smokers' NCDS = 9%, BCS70 = 18%; 'Mainstream' NCDS = 7%, BCS70 = 19%, $p < 0.01$).

3.4. Cluster membership

For both genders, Wald chi-square tests indicated a significant difference in cluster membership across the cohorts. For men and women, a significantly higher proportion ($p < 0.01$) of BCS70 participants (men = 73.9%; women = 76.8%) were members of 'Mainstream' cluster compared to NCDS participants (men = 68.3%; women = 68.8%).

3.5. Sensitivity analysis

Estimates from models using FIML were very similar to estimates using complete cases (see Appendix D).

4. Discussion

Using multi-group LPA, we identified three distinct clusters of HRBs: 'Risky', 'Moderate Smokers' and 'Mainstream'. For both genders, cluster patterns were similar across the two cohorts in relation to smoking, fruit and vegetable consumption, fried food consumption and physical activity. HRBs' of members in the 'Mainstream' cluster tended to be more beneficial to health than the other two clusters (i.e. not smoking, eating fruit and vegetables more frequently, chips and fried food less frequently and being more active), based upon evidence linking these HRBs to mortality (Loef and Walach, 2012; Kvaavik et al., 2010). However, the frequency of sweet food consumption was generally higher in the 'Mainstream' cluster. Moreover, in the later born cohort there was a significant shift in membership towards the 'Mainstream' cluster.

The distribution of alcohol consumption across the three clusters differed by cohort in both genders but was particularly apparent for women. The proportion of BCS70 women drinking above recommended guidelines across the three clusters was almost double than that of NCDS women. At the same time, BCS70 women in the 'Mainstream' cluster consumed sweet foods less frequently than NCDS women.

Our findings add support to previous evidence for HRB clustering (Schuit et al., 2002; Berrigan et al., 2003; Poortinga, 2007; Silva et al., 2013; Laaksonen et al., 2001; van Nieuwenhuijzen et al., 2009; De Vries et al., 2008; Conry et al., 2011; Schneider et al., 2009; Vermeulen-Smit et al., 2015; Leventhal et al., 2014; Tobias et al., 2007; Verger et al., 2009). Research using cluster analysis in large population based samples has identified cluster patterns very similar to ours. Previous research found that people who smoke daily tend to have both lower fruit and vegetable consumption and lower levels of physical activity (Schuit et al., 2002; Laaksonen et al., 2001; De Vries et al., 2008; Conry et al., 2011; Schneider et al., 2009; Vermeulen-Smit et al., 2015; Verger et al., 2009), people who drink alcohol heavily are more likely to smoke (Conry et al., 2011; Schneider et al., 2009; Vermeulen-Smit et al., 2015; Leventhal et al., 2014; Verger et al., 2009) and people who smoke and drink heavily are more likely to consume fried food and less likely to consume sweet snacks (Vermeulen-Smit et al., 2015).

We found that membership of the 'Mainstream' cluster was higher in the BCS70, compared to the NCDS, while membership of the 'Risky' cluster was higher amongst NCDS women compared to BCS70 women. The shift to the 'Mainstream' cluster is beneficial for health in some respects, especially cigarette smoking. Moreover, BCS70 members had higher frequencies of fruit and vegetable consumption, lower frequencies of fried food consumption and were more physically active in this cluster, compared to NCDS members. However, a higher proportion of BCS70 men and women were drinking alcohol above the recommended guidelines (DOH. Sensible drinking: Report of an interdepartmental working group. London: Department of Health, 1995) in the 'Mainstream' cluster and frequency of sweet food consumption tended to be higher in this cluster compared to the other two. Amongst women, sweet food consumption frequency was lower in the 'Mainstream' cluster for BCS70 compared to NCDS members.

Higher membership to the 'Mainstream' cluster in the later born cohort suggests changes in social norms for smoking, alcohol and food consumption. For example, a large proportion of participants in the 'Mainstream' cluster did not smoke cigarettes, corresponding with declines in the prevalence of smoking over the past 50 years (RCP. Fifty years since smoking and health. London: Royal College of Physicians, 2012). Convergence in alcohol consumption of men and women over time in our study is consistent with previous work (Schoon and Parsons, 2003; Elliott et al., 2007; Meng et al., 2014; Keyes et al., 2011) and does not bode well for the 2016 UK government drinking recommendations (DOH, 2016), given that women increased their drinking consumption to meet men's rather than vice versa. Moreover, higher frequency of sweet food consumption in this cluster appears to coincide with global trends of increasing sugar consumption (Singh et al., 2015; WHO, 2015).

A higher proportion of 'Risky' cluster members (NCDS men, BCS70 men and women) reported drinking 'no units' in the previous week. Although members of the 'Risky' cluster drink differently (i.e. not drinking or drinking excessively) they are assigned to the same cluster by sharing other behaviours, particularly smoking. Research investigating these four HRBs suggests smoking to be the most persistent (Paavola et al., 2004) and strongly associated with heavy alcohol consumption (Chiolero et al., 2006; Bien and Burge, 1990; Zacny, 1989; Room, 2004) and alcohol abstainers who have previously drunk alcohol (De Leon et al., 2007).

We found largely consistent cluster patterns of smoking, fruit and vegetable consumption, fried food consumption and physical activity across the two cohorts for both genders, implying that these clusters could be generalised to individuals in mid-adulthood in Britain today. Although other age groups were not explored in this study, HRBs in adulthood were found to be relatively stable (Mulder et al., 1998; Harrington et al., 2014; Telama, 2009; Parsons et al., 2006). This implies that these behaviours could be sustained across adulthood.

4.1. Strengths and limitations of the study

This study maximised the efficiencies of data reduction techniques by treating variables in the study model as continuous or ordered; identifying clusters that may have been missed if variables were dichotomised. LPA allows for the investigation of multiple rather than individual HRBs (McAloney et al., 2013) and provided new insights with existing data by detecting a previously unobserved mixture of three clusters. The study detected cohort differences in HRB clustering according to gender, made possible by the large sample size.

To aid LPA model convergence the mean and variance of smoking in the 'Mainstream' cluster was set at 0 because smoking in this cluster was rare (cigarettes per day: NCDS mean = 0.5; BCS70 mean = 0.3). Sensitivity analysis indicated that this decision did not affect the LPA model estimates (results not shown), with only a small proportion of smokers (NCDS = 7.6%, BCS70 = 5.2%) in this cluster.

It was necessary to use data from ages 30 and 34 in the BCS70, because dietary information was not available at age 34. We assumed dietary habits remained relatively stable during this period based upon empirical evidence (Mulder et al., 1998; Parsons et al., 2006). However, this approach left 663 individuals (7%) with incomplete data. Men had more missing data on diet at age 30 than women ($p < 0.001$). However, similar estimates were found in sensitivity analysis comparing models using FIML and those using complete cases.

During the twelve years separating the two cohorts the average serving size of spirits and wine has increased (Stead et al., 2013), potentially underestimating alcohol consumption amongst BCS70 participants. We found a higher proportion of participants in the BCS70 drinking above recommended limits compared to the NCDS. Therefore, correcting for a potential underestimation of alcohol consumption in the BCS70 would not change the direction of our findings.

This study relies on self-reported measures of HRB which can be biased (Conry et al., 2011; Schneider et al., 2009; Heroux et al., 2012; Schnuerer et al., 2015). However, both cohorts collected data on HRB variables using well-structured questionnaires and in the BCS70 all interviews were assisted with a computer, reducing interviewer error. Although HRB measures were not identical across the cohorts using data from two purposefully similar birth cohort studies allowed for a valid comparison (Ekinsmyth et al., 1992) and reduced the likelihood of cohort and gender differences due to differential measurement. Furthermore, differential measurement would suggest a uniformed bias, instead we saw larger cohort differences amongst women, indicating other contextual factors are at play.

Social desirability bias is a possibility and may explain, to some extent, the size of the 'Mainstream' cluster. However, there are similarities between the prevalence of current smoking in both cohorts (NCDS 1991 = 32%, BCS70 2004 = 24%) and that reported by the Health and Social Care Information Centre (HSCIC), for persons aged 35–49 during the same time periods (1990 = 34%, 2004 = 29%) (HSCIC. Statistics on Smoking, England, 2014).

Also interpretation of cohort differences in cluster membership requires caution due to differences in cluster patterns e.g. alcohol consumption in the 'Risky' cluster for NCDS women resembled that of the 'Moderate Smokers' cluster for BCS70 women. This indicates partial measurement invariance (Collins and Lanza, 2010) i.e. the clusters cannot be interpreted the same way across cohorts. However, the measurement invariance analysis suggested equivalence of the 'Moderate Smokers' and 'Mainstream' clusters across the two cohorts (see Appendix B). This work identified consistent cluster patterns for smoking, fruit and vegetable consumption, fried food consumption and physical activity. Therefore, we think that the 'Moderate Smokers' and 'Mainstream' cluster patterns, are likely to be similar amongst individuals in mid-adulthood in Britain more generally.

4.2. Implications

Whilst the number of studies evaluating the efficacy of multiple HRB interventions is limited (King et al., 2015) and comparability of the study samples with ours is questionable, they suggest that interventions addressing multiple HRBs simultaneously may be more successful (Ashra NBS et al., 2015; Hale et al., 2014; Steptoe, 2007; Hyman et al., 2007; Goldstein et al., 2004; Nigg et al., 2002) and cost effective (Prochaska et al., 2008) than targeting HRBs independently. Our findings reinforce this evidence, in our study smokers consumed fruits and vegetables infrequently, but consumed fried food frequently. Any interventions to quit smoking could employ person-centred strategies to improve diets. Our findings showed that sweet food consumption in the 'Mainstream' cluster tended to be higher than the other two clusters. This cluster pattern along with the shift in membership to the 'Mainstream' cluster amongst those born in 1970, could partially explain increases in overweight and obesity rates observed in the later born cohort (Sullivan et al., 2013). We found significant increases in alcohol consumption and decreases in frequency of sweet food consumption for female members of the 'Mainstream' cluster born in 1970 compared to 1958. This may reflect a replacement of sugar intake with alcohol use amongst women in the later born cohort (Colditz et al., 1991) given that alcoholic beverages, particularly types popular amongst women (i.e. cocktails), are often calorie dense (Sayon-Orea et al., 2011).

Further analysis to identify common predictors of cluster membership across cohorts could strengthen our case that the 'Moderate Smokers' and 'Mainstream' clusters are similar amongst individuals in mid-adulthood today and assist in developing person-centred interventions.

5. Conclusions

In 2 representative British cohorts we found three clusters of HRBs labelled: 'Risky', 'Moderate Smokers' and 'Mainstream'. Our findings suggest largely consistent HRB cluster patterns across cohort and gender groups which could be found amongst people in mid-adulthood in Britain today. This new evidence of HRB clustering across time and by gender provides a person-centred understanding that can inform population level interventions to improve HRBs.

Conflict of interest statement

The authors declare that there are no conflicts of interest.

Acknowledgments

The authors would like to acknowledge the Economic and Social Research Council (grant ES/J500185/1; ES/J019119/1) and the European Research Council (grant ERC-StG-2012-309337) for funding this research. The authors would also like to acknowledge Professor Gopal Netuveli and Dr. Paul Watts from the University of East London for their feedback on a preliminary version of the manuscript and Professor Ingrid Schoon from the Institute of Education, University College London for her suggestions on health-related behaviour measurement.

Appendix A. Cohort member questionnaire item wording in the NCDS at age 33 (all variables) and BCS70 at age 30 (diet) and 34 (smoking, physical activity, alcohol consumption) and cohort harmonisation

Variable	Question NCDS	Response category NCDS	Question BCS70	Response category BCS70	Cohort harmonisation
Smoking (cigarettes smoked per day)	1) Do you smoke cigarettes at all nowadays? 2) How many cigarettes a day do you usually smoke?	Yes No	1) Now some questions about smoking. Would you say that: 2) How many cigarettes a day do you usually smoke?	a) You've never smoked cigarettes? b) You used to smoke cigarettes but don't at all now? c) You now smoke cigarettes occasionally but not every day? d) You smoke cigarettes every day?	Question 1 NCDS No = 0 Question 1 BCS70 a/b/c = 0 Question 2 NCDS/BCS70 response = cigarettes smoked per day (range 0–80) NOTE 1: BCS70 participants who answered a/b/c for question 1 were not asked question 2. NOTE 2: Sensitivity analysis in the BCS70 found combining 'occasional' smokers with daily smokers, rather than non-smokers, did not influence Latent Profile Analysis model estimates.

Appendix A (*continued*)

Variable	Question NCDS	Response category NCDS	Question BCS70	Response category BCS70	Cohort harmonisation
Frequency of fruit and vegetable consumption	1) How often do you eat fresh fruit in summer? 2) How often do you eat salads or raw vegetables in winter?	a) More than once a day b) Once a day c) 3–6 days a week d) 1 or 2 days a week e) Less than 1 day a week f) Never	1) How often do you eat fresh fruit? 2) How often do you eat salads or raw vegetables? 3) How often do you eat cooked veg?	a) More than once a day b) Once a day c) 3–6 days a week d) 1 or 2 days a week e) Less than 1 day a week f) Occasionally g) Never	Questions 1 and 2 NCDS f = 0, e = 1, d = 2, c = 3, b = 4, a = 5 Questions 1 and 2 BCS70 g = 0, f = 1, e = 1, d = 2, c = 3, b = 4, a = 5 Q1 diet score + Q2 diet score = FV diet score (1–10)
Frequency of chips and fried food consumption	1) How often do you eat chips? 2) How often do you eat fried food not including chips?	a) More than once a day b) Once a day c) 3–6 days a week d) 1 or 2 days a week e) Less than 1 day a week f) Never	1) How often do you eat food fried in vegetable oil such as olive oil or sunflower oil, not counting chips? 2) How often do you eat food fried in hard fat such as lard or butter, not counting chips? 3) How often do you eat chips?	a) More than once a day b) Once a day c) 3–6 days a week d) 1 or 2 days a week e) Less than 1 day a week f) Occasionally g) Never	Questions 1 and 2 NCDS f = 0, e = 1, d = 2, c = 3, b = 4, a = 5 Questions 1, 2 and 3 BCS70 g = 0, f = 1, e = 1, d = 2, c = 3, b = 4, a = 5 Q1 diet score + Q2 diet score = CF diet score (1–10)
frequency of sweets, chocolate, biscuits/cake consumption	1) How often do you eat sweets, chocolates? 2) How often do you eat biscuits?	a) More than once a day b) Once a day c) 3–6 days a week d) 1 or 2 days a week e) Less than 1 day a week f) Never	1) How often do you eat sweets or chocolates? 2) How often do you eat biscuits and cakes of all kinds?	a) More than once a day b) Once a day c) 3–6 days a week d) 1 or 2 days a week e) Less than 1 day a week f) Occasionally g) Never	Questions 1 and 2 NCDS f = 0, e = 1, d = 2, c = 3, b = 4, a = 5 Questions 1 and 2 BCS70 g = 0, f = 1, e = 1, d = 2, c = 3, b = 4, a = 5 Q1 diet score + Q2 diet score = SCBC diet score (1–10)
Frequency of leisure time physical activity	1) Do you regularly take part in any activities on this card (see below) - that is at least once a month for most of the year? Activities listed on card •Any competitive sports •'Keep fit' or aerobics classes •Circuit training •Weight training or other repeated exercises (at home or in the gym) (listed only at age 33 NCDS) •Running or jogging •Swimming •Cycling •Going for walks •Taking part in water sports •Outdoor sports •Dancing •Any other sport or leisure activity that involves physical exercise	Yes No	1) Do you regularly take part in any of the activities on this card (see below), by regularly I mean at least once a month, for most of the year. Activities listed on card •Any competitive sports •'Keep fit' or aerobics classes •Circuit training •Running or jogging •Swimming •Cycling •Going for walks •Taking part in water sports •Outdoor sports •Dancing •Any other sport or leisure activity that involves physical exercise	Yes No	Question 1 NCDS/BCS70 No = 0 Question 2 NCDS/BCS70 f = 0, e = 0, d = 1, c = 2, b = 3, a = 4 Q1 and Q2 combined = Frequency of leisure time physical activity with 4 categories. '≤3 times a month' 'Once a week' '2–3 days a week' '4–7 days a week'. NOTE: 6 response categories collapsed into 4 due to sparseness.
	2) How often do you take part in any activity of this type?	a) Every day b) 4–5 days a week c) 2–3 days a week d) once a week e) 2–3 times a month f) Less often	2) How often do you take part in any activity of this type?	a) Every day b) 4–5 days a week c) 2–3 days a week d) once a week e) 2–3 times a month f) Less often	
Alcohol consumption (units consumed in the previous week)	1) How often do you have an alcohol drink of any kind? 2) In the last 7 days, that is not counting today by starting from last [name present day of week], how much beer/stout/lager/ale/cider have you had? 3) In the last 7 days how many measures of spirits or liqueurs have you had? 4) In the last 7 days how many glasses of wine have you had?	a) Most days b) 1,2,3 times a week c) 1,2,3 times a month d) Less often or only on special occasions e) Never	1) How often do you have an alcoholic drink of any kind? Would you say you had a drink 2) In the last 7 days that is now counting today but starting from last [^Day7Ago], how much beer/stout/lager/ale/cider have you had? 3) In the last 7 days how many measures of spirits or liqueurs have you had, like Gin, Whiskey, Rum, Brandy, Vodka or Advocate?	a) On most days b) 2–3 days a week c) Once a week d) 2–3 times a month e) Less often or only on special occasions f) Never now a days g) Have you never had an alcohol drink?	Question 1 NCDS e = 0, d = 0, c = 1, b = 1, a = 1 Question 1 BCS70 g = 0, f = 0, e = 0, d = 1, c = 1, b = 1, a = 1 Question 2 NCDS measured in pints. Values converted to units (multiplied by 2). Question 3, 4 and 5 NCDS measured in units. Question 2, 3, 4, 5, 6, and 7 BCS70 measured in units. Alcohol consumption NCDS='no units' if Q1,Q2,Q3, Q4,Q5 = 0 Alcohol consumption bcs70='no units' if Q1,Q2,Q3,

(continued on next page)

Appendix A (*continued*)

Variable	Question NCDS	Response category NCDS	Question BCS70	Response category BCS70	Cohort harmonisation
					Q4,Q5,Q6,Q7 = 0 Alcohol consumption NCDS='within limits' if Q1 = 1 and Q2 + Q3 + Q4 + Q5 = 1 to 21 for men, 1 to 14 for women. Alcohol consumption BCS70='within limits' if Q1 = 1 and Q2 + Q3 + Q4 + Q5 + Q6 + Q7 = ≥22 for men, ≥15 for women. NOTE 1: One outlier in BCS70 (reported 280 units) coded as missing. NOTE 2: Sensitivity analysis found that including values from Q6 and Q7 in BCS70 did not inflate alcohol consumption in this cohort (only a small number of participants drank alcopops or other drinks). NOTE 3: Sensitivity analysis found including 'never drinkers' as a separate category did not influence Latent Profile Analysis model estimates.
	5) In the last 7 days how many glasses of martini/vermouth/sherry or similar drinks have you had?		4) In the last 7 days how many glasses of wine have you had? 5) In the last 7 days how many glasses of martini/ vermouth/sherry /port or similar drinks have you had? 6) In the last 7 days how many bottles of alcopops have you had? 7) In the last 7 days have you had any other alcohol drinks?		

Appendix B. Measurement invariance analysis

The measurement invariance analysis was conducted in five parts:

Firstly, we established within each subgroup (separated according to gender and cohort) that the three cluster model was preferred over a 2 and 4 cluster model. The results of this analysis (described fully in Appendix C) suggested that in all subgroups the 3 cluster model was preferred, indicating some equivalence of the clusters across the cohorts.

Secondly, chi-square difference tests using the loglikelihood from multi-group LPA models were conducted; model one allowed values of the observed indicator variables to be free across the cohorts within each gender; model two fixed the observed indicator values to be the same across the cohorts within each gender. We found a significant p-value in all instances, indicating that the fit of the model with fixed (constrained) parameters was substantially worse. Therefore the measurement models for the two cohorts differed significantly and full measurement invariance did not hold (Finch, 2015). The results of this analysis are presented in Tables B1 and B2.

However, scholars argue that such hypothesis testing can be difficult in LPA because models can become very sparse (containing cells with few participants) meaning that the loglikelihood chi-square ratio distribution is not adequately approximated by the chi-square (76). Additionally, with large sample sizes ($N > 2000$) the loglikelihood chi-square difference test can detect minor differences indicating non-invariance for small differences across groups (77).

Therefore, the third step was to conduct cross-validation analysis (Collins and Lanza, 2010), which estimates a LPA model based on group 1 (training dataset) and applies these parameters to group 2 (validation dataset) and vice versa. This is done to determine whether the model calibrated in the training dataset has an acceptable model fit in the validation dataset. Assessing whether individuals remain within the same cluster in the validated and calibrated models is also an indication of membership stability to a particular cluster (i.e. classification certainty) (Collins and Lanza, 2010). Models that cross-validate well indicate that the nature of the latent clusters are similar across the groups and that measurement invariance holds (76). Based on model fit indices and entropy, we found that models calibrated in each subgroup (separated by cohort and gender) cross-validated well when applied to data from the same gender in the other cohort. Tables B3 and B4 outline the results of this analysis.

Table B1
Estimates from multi-group models with and without cluster patterns and membership constrained to be equal (men).
Data: two British birth cohort studies, the National Child Development Study (NCDS) at age 33 (1991), the British Cohort Study (BCS70) at age 30 (2000) and 34 (2004).

Men FIML multi-group 3 cluster model	Loglikelihood	Scaling correction factor	Number of parameters	Chi-square difference (p value)
Cluster patterns free	−73,256.573	1.2612	83	533.69 (<0.001)
Cluster patterns fixed	−73,841.447	1.0215	66	
Cluster membership free	−73,256.573	1.2612	83	32.65 (<0.001)
Cluster membership fixed	−73,276.564	1.2621	81	

Table B2
Estimates from multi-group models with and without cluster patterns and membership constrained to be equal (women).
Data: two British birth cohort studies, the National Child Development Study (NCDS) at age 33 (1991), the British Cohort Study (BCS70) at age 30 (2000) and 34 (2004).

Women FIML multi-group 3 cluster model	Loglikelihood	Scaling correction factor	Number of parameters	Chi-square difference (p value)
Cluster patterns free	−75,233.802	1.1511	83	870.54 (<0.001)
Cluster patterns fixed	−75,896.731	1.0553	66	
Cluster membership free	−75,233.802	1.1511	83	43.80 (<0.001)
Cluster membership fixed	−75,279.763	1.1277	81	

Table B3
Estimates from cross-validation models (men).
Data: Two British birth cohort studies, the National Child Development Study (NCDS) at age 33 (1991), the British Cohort Study (BCS70) at age 30 (2000) and 34 (2004).

Men 3 cluster FIML models[a]	Loglikelihood	Entropy	Adjusted BIC	Lo–Mendell-Rubin LRT *p* value[b]
NCDS Men calibrated	−38,559.333	0.981	77,341.679	0.04
NCDS Men validated	−40,265.503	0.950	80,531.007	<0.001
BCS70 Men calibrated	−26,838.585	0.980	53,888.898	<0.001
BCS70 Men validated	−27,698.580	0.992	55,397.160	<0.001

[a] Calibrated models = model parameters estimated in this gender and cohort. Validated models = gender and cohort in which the model is validated using saved model parameter estimates from the model calibrated in the same gender but opposite cohort data.
[b] Tests the null hypotheses that the addition of a fourth cluster does not improve model fit.

Table B4
Estimates from cross-validation models (women).
Data: Two British birth cohort studies, the National Child Development Study (NCDS) at age 33 (1991), the British Cohort Study (BCS70) at age 30 (2000) and 34 (2004).

Women 3 cluster FIML models[a]	Loglikelihood	Entropy	Adjusted BIC	Lo–Mendell-Rubin LRT *p* value[b]
NCDS women calibrated	−38,558.954	0.901	77,342.313	<0.001
NCDS women validated	−40,604.876	0.907	81,209.751	<0.001
BCS70 women calibrated	−28,595.963	0.943	57,408.650	0.02
BCS70 women validated	−29,787.017	0.933	59,574.034	>0.05

[a] Calibrated models = model parameters estimated in this gender and cohort. Validated models = gender and cohort in which the model is validated using saved model parameter estimates from the model calibrated in the same gender but opposite cohort data.
[b] Tests the null hypotheses that the addition of a fourth cluster does not improve model fit.

Table B5
The stability of cluster membership in calibrated and validated models (NCDS men).
Data: The National Child Development Study (NCDS) at age 33 (1991).

NCDS Men	'Risky' cluster validated n (%)	'Moderate Smokers' cluster validated n (%)	'Mainstream' cluster validated n (%)	Total
'Risky' cluster calibrated	26 (8.1)	56 (3.9)	0	82
'Moderate Smokers' cluster calibrated	288 (89.2)	1398 (96.2)	0	1686
'Mainstream' cluster calibrated	9 (2.8)	0	3809 (100.0)	3818
Total	323 (100.0)	1454 (100.0)	3809 (100.0)	5586

Table B6
The stability of cluster membership in calibrated and validated models (BCS70 men).
Data: The British Cohort Study (BCS70) at age 30 (2000) and 34 (2004).

BCS70 Men	'Risky' cluster validated n (%)	'Moderate Smokers' cluster validated n (%)	'Mainstream' cluster validated n (%)	Total
'Risky' cluster calibrated	11 (44.0)	65 (5.5)	3 (0.1)	79
'Moderate Smokers' cluster calibrated	14 (56.0)	1110 (94.5)	0	1124
'Mainstream' cluster calibrated	0	0	3410 (99.9)	3410
Total	25 (100.0)	1175 (100.0)	3413 (100.0)	4613

Table B7
The stability of cluster membership in calibrated and validated models (NCDS women).
Data: The National Child Development Study (NCDS) at age 33 (1991).

NCDS Women	'Risky' cluster validated n (%)	'Moderate Smokers' cluster validated n (%)	'Mainstream' cluster validated n (%)	Total
'Risky' cluster calibrated	444 (78.6)	70 (5.6)	1 (0.1)	515
'Moderate Smokers' cluster calibrated	121 (21.4)	1171 (94.4)	0	1292
'Mainstream' cluster calibrated	0	0	3980 (99.9)	3980
Total	565 (100.0)	1241 (100.0)	3981 (100.0)	5787

Table B8
The stability of cluster membership in calibrated and validated models (BCS70 women).
Data: The British Cohort Study (BCS70) at age 30 (2000) and 34 (2004).

BCS70 Women	'Risky' cluster validated n (%)	'Moderate Smokers' cluster validated n (%)	'Mainstream' cluster validated n (%)	Total
'Risky' cluster calibrated	136 (76.8)	46 (4.7)	1 (0.1)	183
'Moderate Smokers' cluster calibrated	41 (23.2)	943 (95.4)	0	984
'Mainstream' cluster calibrated	0	0	3866 (99.9)	3866
Total	177 (100.0)	989 (100.0)	3867 (100.0)	5033

The stability of cluster membership in the validated and calibrated models was deemed to be excellent for the 'Mainstream' and 'Moderate Smokers' clusters with ≥95% of individuals being assigned to the same cluster in the calibrated and validated models. Cluster classification was also deemed to be good for the 'Risky' cluster amongst women given that ≥77% remained in the same cluster in the calibrated and validated models. However, there appeared to be classification uncertainty for the 'Risky' cluster amongst men. This could, in part, be due to the small number of participants assigned to this cluster in the calibrated models. The results of this analysis are presented in Tables B5–B8.

Fourthly, alongside these assessments of measurement invariance, scholars suggest that researchers consider if the nature of the clusters differ according to groups by examining differences in cluster membership (i.e. prevalence) and cluster patterns (i.e. item means and probabilities) (Collins and Lanza, 2010). Evidenced statistically by conducting chi-square Wald tests to identify if levels of the health-related behaviours differed significantly both across the clusters within each cohort and within each cluster across cohorts. These cluster patterns and the results of the chi-square Wald test are presented in the main body of the manuscript (see Tables 2 and 3).

Our interpretation was that cluster patterns across the cohorts were similar, except for alcohol consumption, particularly amongst women and suggested partial measurement invariance across cohorts and genders. This is concurrent with other research that has found convergence between men and women in alcohol consumption in the later born cohort (Elliott et al., 2007). We attempted to run a LPA model which pooled the cohort data (increasing statistical power) whilst accounting for alcohol differences in the model. However, due to increased complexity and data sparseness (given the size of the smallest cluster) the pooled LPA model would not converge and therefore the decision was taken to present LPA results separately according to cohort and gender in the manuscript.

Table B9
Indicator variable specific entropy (Men).
Data: Two British birth cohort studies, the National Child Development Study (NCDS) at age 33 (1991), the British Cohort Study (BCS70) at age 30 (2000) and 34 (2004).

Indicator variable	Indicator variable specific entropy	
	NCDS Men	BCS70 Men
Smoking	0.97	0.93
Alcohol	0.38	0.43
Fruit and vegetables	0.39	0.42
Fried food	0.39	0.47
Sweet food	0.38	0.41
Physical activity	0.38	0.42

Table B10
Indicator variable specific entropy (Women).
Data: Two British birth cohort studies, the National Child Development Study (NCDS) at age 33 (1991), the British Cohort Study (BCS70) at age 30 (2000) and 34 (2004).

Indicator variable	Indicator variable specific entropy	
	NCDS Women	BCS70 Women
Smoking	0.84	0.91
Alcohol	0.26	0.41
Fruit and vegetables	0.32	0.42
Fried food	0.30	0.43
Sweet food	0.27	0.41
Physical activity	0.28	0.41

Finally, Mplus version 7.3 (Muthen, 2014) provides indicator variable specific entropy which measures the extent to which each observed variable in the model determines the unobserved (latent) variable. An entropy of below <0.2 suggests that the observed health-related behaviour variable contributes little in defining the latent health-related behaviour clusters (78). We found in all LPA models (separated by cohort and gender) that the four health-related behaviours had entropy above 0.2 and that entropy values were similar across the cohorts, but differed more for women than men. This provided further evidence of cluster equivalence across the cohorts because the contribution of individual health-related behaviours to the latent variable was similar within each gender group across the clusters. Although, it should be noted that the observed difference in cluster patterns for alcohol consumption amongst women combined with information gained from indicator variable entropy (i.e. there is a greater contribution of alcohol to the formation of health-related behaviour clusters amongst BCS70 women than NCDS women), suggests partial measurement invariance for alcohol consumption amongst women across the cohorts. Tables B9 and B10 outline the indicator variable specific entropy for each subgroup.

In summary, on the basis of the above analysis we conclude that the nature of the clusters is similar within each gender group across the two cohorts. However, the results do suggest that the clusters may not be equivalent for alcohol consumption amongst women and that the stability of membership to the 'Risky' cluster amongst men in the two cohorts is questionable.

References

21. Elliott J, Dodgeon B, Elliott J. A Descriptive Analysis of the Drinking Behaviour of the 1958 Cohort at Age 33 and the 1970 Cohort at Age 34. London: Centre for Longitudinal Studies, 2007. Available from: http://www.cls.ioe.ac.uk/library-media/documents/CLS_WP_2007_3 (WHO, 2014).pdf.

34. Collins LM, Lanza ST. Latent class and latent transition analysis: With applications in the social, behavioural, and health sciences: John Wiley & Sons; 2010.

35. Muthen M. Mplus: Version 7. Los Angeles, CA Muthen & Muthen; 2014.

40. Finch H. A Comparison of Statistics for Assessing Model Invariance in Latent Class Analysis. Open Journal of Statistics. 2015;5(03):191.

76. Collins LM, Graham JW, Long JD, Hansen WB. Crossvalidation of latent class models of early substance use onset. Multivariate Behavioural Research. 1994;29 (2):165–83.

77. Meade AW, Lautenschlager GJ. A Monte-Carlo Study of Confirmatory Factor Analytic Tests of Measurement Equivalence/Invariance. Structural Equation Modelling: A Multidisciplinary Journal. 2004;11 (1):60–72. DOI: 10.1207/S15328007SEM1101_5.

78. Asparouhov T, Muthén B. Variable-Specific Entropy Contribution. Los Angeles: Muthén & Muthén, 2014. Available from: https://www.statmodel.com/download/UnivariateEntropy.pdf.

Appendix C. Latent Profile Analysis model fit indices and cluster power

The selection of a 3 cluster multi-group LPA model was based upon model fit indices and minimum cluster size criterion presented in Table B1. The purpose of the minimum cluster size criterion was to ensure adequate statistical power for further analysis. Adequate cluster size was determined by detecting a 'small' difference using Cohen's effect size of 20% (79) between two independent proportions (0.1, 0.3), with 80% power and significance level of 0.05. Sample size calculations were conducted in Stata version 13 (80) using the 'power two proportions' command.

Table B1
Goodness of fit indices for Latent Profile Analysis (LPA) models, stratified by cohort for each gender. Data: Two British birth cohort studies, the National Child Development Study (NCDS) at age 33 (1991), the British Cohort Study (BCS70) at age 30 (2000) and 34 (2004).

NCDS Men	Loglikelihood	aBIC	LMR	Entropy	Smallest cluster size (n)	MMC (n = 124)[a]
2 cluster	39,034.135	78,220.878	<0.001	0.990	1768 (31.7%)	Yes
3 cluster	38,917.438	78,058.340	0.03	0.978	82 (1.5%)	No
4 cluster	38,850.141	77,994.599	0.06	0.917	80 (1.4%)	No
BCS70 Men	**Loglikelihood**	**aBIC**	**LMR**	**Entropy**	**Smallest Cluster Size (n)**	**MMC (n = 124)[a]**
2 cluster	27,440.492	55,028.237	<0.001	0.993	1200 (26.0%)	Yes
3 cluster	27,316.210	54,848.039	0.04	0.975	79 (1.7%)	No
4 cluster	27,214.809	54,713.604	0.003	0.981	24 (0.5%)	No
NCDS Women	**Loglikelihood**	**aBIC**	**LMR**	**Entropy**	**Smallest Cluster Size (n)**	**MMC (n = 124)[a]**
2 cluster	39,075.405	78,304.409	<0.001	0.991	1808 (31.2%)	Yes
3 cluster	38,915.657	78,056.225	<0.001	0.899	515 (8.9%)	Yes
4 cluster	38,849.292	77,994.811	0.01	0.905	59 (1.0%)	No
BS70 Women	**Loglikelihood**	**aBIC**	**LMR**	**Entropy**	**Smallest Cluster Size (n)**	**MMC (n = 124)[a]**
2 cluster	28,922.370	57,994.432	<0.001	0.993	1166 (23.2%)	Yes
3 cluster	28,844.586	57,908.362	0.02	0.940	183 (3.6%)	Yes
4 cluster	28,789.143	57,866.976	0.09	0.951	33 (0.7%)	No

• Note: aBIC = adjusted Bayesian Information Criteron; LMR = Lo–Mendell–Rubin Likelihood Ratio Test *p* value. MMC = Meets Minimum Cluster Size Criterion.
[a] Adequate cluster size determined by detecting a 20% difference in two independent proportions (0.1, 0.3), with 80% power and significance level of 0.05.

79. Cohen J. A power primer. Psychological bulletin. 1992;112(1):155. DOI:http://dx.doi.org/10.1037/0033-2909.112.1.155
80. StataCorp. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP; 2013.

Appendix D. Latent Profile Analysis model estimates excluding individuals with incomplete information on health-related behaviour variables

Table D1
Estimated means and item response probabilities of 3 cluster multiple-group Latent Profile Analysis (LPA) model, using complete cases, for men.
Data: two British birth cohort studies, the National Child Development Study (NCDS) at age 33 (1991), the British Cohort Study (BCS70) at age 30 (2000) and 34 (2004).

	NCDS men total N = 5525 (100%)			BCS70 men total N = 4195 (100%)		
	Cluster 1 'Risky' n = 79 (1.4%)≠	Cluster 2 'Moderate Smokers' n = 1677 (30.4%)≠	Cluster 3 'Mainstream' n = 3769 (68.2%)≠	Cluster 1 'Risky' n = 78 (1.9%)≠	Cluster 2 'Moderate Smokers' n = 1003 (23.9%)≠	Cluster 3 'Mainstream' n = 3114 (74.2%)≠
	Mean (S.E)	Mean (S.E)	Mean (S.E)	Mean (S.E)	Mean (S.E)	Mean (S.E)
Number of cigarettes smoked per day	41.24 (3.88)*†	17.23 (0.32)*†	0*	18.14 (2.15)†	15.60 (0.26)†	0
Frequency of fruit and vegetable consumption	2.62 (0.42)*	3.95 (0.05)*†	4.64 (0.03)*†	3.77 (0.27)*	4.29 (0.07)*†	5.10 (0.04)*†
Frequency of fried food consumption	4.69 (0.51)*†	3.99 (0.05)*†	3.36 (0.02)*†	6.50 (0.27)*†	3.01 (0.04)*†	2.86 (0.02)*†
Frequency of sweet food consumption	3.56 (0.49)*†	4.18 (0.06)*	4.71 (0.04)*	5.29 (0.49)*†	4.34 (0.08)*	4.59 (0.04)*
	Item Response probability (S.E)	Item Response probability (S.E)	Item Response probability (S.E)	Item Response probability (S.E)	Item Response probability (S.E)	Item Response probability (S.E)
Frequency of leisure time physical activity						
≤3 times a month	0.60 (0.07)*	0.39 (0.01)*	0.28 (0.01)*†	0.49 (0.07)*	0.40 (0.02)*	0.27 (0.01)*†
Once a week	0.13 (0.06)	0.21 (0.01)	0.21 (0.01)	0.09 (0.04)	0.18 (0.01)	0.18 (0.01)
2–3 days a week	0.14 (0.05)	0.19 (0.01)	0.25 (0.01)	0.21 (0.05)	0.20 (0.01)	0.30 (0.01)
4–7 days a week	0.13 (0.05)	0.21 (0.01)	0.26 (0.01)	0.21 (0.05)	0.22 (0.01)	0.26 (0.01)
Alcohol units consumed in the previous week						
No units	0.25 (0.08)*	0.14 (0.01)*†	0.13 (0.01)*†	0.27 (0.06)*	0.14 (0.01)*†	0.12 (0.01)*†
Within limits (≤14 units women, ≤21 units men)	0.23 (0.06)	0.50 (0.01)	0.63 (0.01)	0.33 (0.06)	0.44 (0.02)	0.61 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.52 (0.09)	0.36 (0.01)	0.24 (0.01)	0.40 (0.08)	0.43 (0.02)	0.27 (0.01)

Note: * = cluster means and response probabilities are significantly different (p ≤ 0.05) across the three clusters within each cohort. † = cluster means and response probabilities are significantly different (p ≤ 0.01) across the cohorts. ≠ = cluster membership is significantly different (p ≤ 0.01) across the cohorts. Estimated using the Wald chi-square test.

Table D2
Estimated means and item response probabilities of 3 cluster multiple-group Latent Profile Analysis (LPA) model, using complete cases, for women.
Data: Two British birth cohort studies, the National Child Development Study (NCDS) at age 33 (1991), the British Cohort Study (BCS70) at age 30 (2000) and 34 (2004).

	NCDS Women Total N = 5716 (100%)			BCS70 Women Total N = 4739 (100%)		
	Cluster 1 'Risky' n = 507 (8.7%)≠	Cluster 2 'Moderate Smokers' n = 1290 (22.6%)≠	Cluster 3 'Mainstream' n = 3919 (68.7%)≠	Cluster 1 'Risky' n = 177 (3.7%)≠	Cluster 2 'Moderate Smokers' n = 908 (19.2%)≠	Cluster 3 'Mainstream' n = 3654 (77.1%)≠
	Mean (S.E)	Mean (S.E)	Mean (S.E)	Mean (S.E)	Mean (S.E)	Mean (S.E)
Number of cigarettes smoked per day	21.01 (1.01)*	14.09 (0.31)*	0	19.27 (1.86)	12.29 (0.39)	0
Frequency of fruit and vegetable consumption	3.38 (0.15)*†	5.56 (0.14)*†	5.79 (0.03)*†	3.66 (0.16)*†	5.42 (0.20)*†	5.97 (0.04)*†
Frequency of fried food consumption	4.03 (0.16)*†	2.70 (0.08)*	2.55 (0.02)*†	3.35 (0.29)*†	2.32 (0.07)*	2.36 (0.02)*†
Frequency of sweet food consumption	3.75 (0.25)*†	4.41 (0.10)*	4.85 (0.04)*†	3.68 (0.27)*†	4.51 (0.12)*	4.60 (0.04)*†
	Item Response probability (S.E)	Item Response probability (S.E)	Item Response probability (S.E)	Item Response probability (S.E)	Item Response probability (S.E)	Item Response probability (S.E)
Frequency of leisure time physical activity						
≤3 times a month	0.62 (0.03)*	0.29 (0.03)*†	0.27 (0.01)*†	0.55 (0.06)*	0.30 (0.03)*†	0.25 (0.01)*†
Once a week	0.16 (0.02)	0.21 (0.01)	0.24 (0.01)	0.08 (0.03)	0.15 (0.01)	0.17 (0.01)
2–3 days a week	0.07 (0.02)	0.20 (0.01)	0.21 (0.01)	0.07 (0.03)	0.23 (0.02)	0.28 (0.01)
4–7 days a week	0.15 (0.02)	0.31 (0.01)	0.28 (0.01)	0.30 (0.05)	0.32 (0.02)	0.31 (0.01)
Alcohol units consumed in the previous week						
No units	0.27 (0.03)*†	0.30 (0.01)*†	0.29 (0.01)*†	0.40 (0.05)*†	0.28 (0.02)*†	0.24 (0.01)*†
Within limits (≤14 units women, ≤21 units men)	0.54 (0.03)	0.61 (0.01)	0.65 (0.01)	0.28 (0.08)	0.53 (0.02)	0.63 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.19 (0.03)	0.09 (0.01)	0.07 (0.01)	0.32 (0.08)	0.18 (0.02)	0.13 (0.01)

Note: * = cluster means and response probabilities are significantly different (p ≤ 0.05) across the three clusters within each cohort. † = cluster means and response probabilities are significantly different (p ≤ 0.01) across the cohorts. ≠ = cluster membership is significantly different (p ≤ 0.01) across the cohorts. Estimated using the Wald chi-square test.

References

Ashra NBS, Rebecca., Carter, Patrice, Davies, Melanie, J., Dunkley, Alison., Gillies, Claire, Greaves, Colin., Khunti, Kamlesh., Sutton, Sarah., Yates, Thomas, Youssef, Dalia., Gray, Laura J., 2015. A Systematic Review and Meta-analysis Assessing the Effectiveness of Pragmatic Lifestyle Interventions for the Prevention of Type 2 Diabetes Mellitus in Routine Practice. Public Health England, London (Available from: https://www.gov.uk/government/publications/diabetes-prevention-programmes-evidence-review).

Berrigan, D., Dodd, K., Troiano, R.P., Krebs-Smith, S.M., Barbash, R.B., 2003. Patterns of health behavior in U.S. adults. Prev. Med. 36 (5), 615–623. http://dx.doi.org/10.1016/S0091-7435(02)00067-1.

Bien, T.H., Burge, R., 1990. Smoking and drinking: a review of the literature. Subst. Use Misuse 25 (12), 1429–1454.

Buck, D., Frosini, F., 2012. Clustering of Unhealthy Behaviours Over Time: Implications for Policy and Practice. The Kings Fund, London (Available from: http://www.kingsfund.org.uk/publications/clustering-unhealthy-behaviours-over-time).

Chiolero, A., Wietlisbach, V., Ruffieux, C., Paccaud, F., Cornuz, J., 2006. Clustering of risk behaviors with cigarette consumption: A population-based survey. Prev. Med. 42 (5), 348–353. http://dx.doi.org/10.1016/j.ypmed.2006.01.011.

Colditz, G.A., Giovannucci, E., Rimm, E.B., Stampfer, M.J., Rosner, B., Speizer, F.E., et al., 1991. Alcohol intake in relation to diet and obesity in women and men. Am. J. Clin. Nutr. 54 (1), 49–55 (http://www.ncbi.nlm.nih.gov/pubmed/2058587. PMCID: 2206044).

Collins, L.M., Lanza, S.T., 2010. Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences. John Wiley & Sons.

Conry, M.C., Morgan, K., Curry, P., McGee, H., Harrington, J., Ward, M., et al., 2011. The clustering of health behaviours in Ireland and their relationship with mental health, self-rated health and quality of life. BMC Public Health 11 (1), 692. http://dx.doi.org/10.1186/1471-2458-11-692.

De Leon, J., Rendon, D.M., Baca-Garcia, E., Aizpuru, F., Gonzalez-Pinto, A., Anitua, C., et al., 2007. Association between smoking and alcohol use in the general population: stable and unstable odds ratios across two years in two different countries. Alcohol Alcohol. 42 (3), 252–257. http://dx.doi.org/10.1093/alcalc/agm029.

De Vries, H., van't Riet, J., Spigt, M., Metsemakers, J., van den Akker, M., Vermunt, J.K., et al., 2008. Clusters of lifestyle behaviors: results from the Dutch SMILE study. Prev. Med. 46 (3), 203–208. http://dx.doi.org/10.1016/j.ypmed.2007.08.005.

DOH, 1995. Sensible Drinking: Report of an Inter-departmental Working Group. Department of Health, London (Available from: http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/publicationsandstatistics/publications/publicationspolicyandguidance/dh_4084701).

DOH, 2016. UK Chief Medical Officers' Alcohol Guidelines Review: Summary of the Proposed New Guidelines. Department of Health, London (Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/489795/summary.pdf).

Ekinsmyth, C., Bynner, J., Montgomery, S., Shepherd, P., 1992. An Integrated Approach to the Design and Analysis of the 1970 British Cohort Study (BCS70) and the National Child Development Study (NCDS). Centre for Longitudinal Studies, London (Available from: http://www.cls.ioe.ac.uk/library-media%5Cdocuments%5Cicwp1.pdf).

Elliott, J., Shepherd, P., 2006. Cohort profile: 1970 British Birth Cohort (BCS70). Int. J. Epidemiol. 35 (4), 836–843. http://dx.doi.org/10.1093/ije/dyl174.

Elliott, J., Dodgeon, B., Elliott, J., 2007. A Descriptive Analysis of the Drinking Behaviour of the 1958 Cohort at Age 33 and the 1970 Cohort at Age 34. Centre for Longitudinal Studies, London (Available from: http://www.cls.ioe.ac.uk/library-media/documents/CLS_WP_2007_3(1).pdf).

Enders, C.K., 2010. Applied Missing Data Analysis. Guilford Press.

Finch, H., 2015. A comparison of statistics for assessing model invariance in latent class analysis. Open J. Stat. 5 (03), 191.

Goldstein, M.G., Whitlock, E.P., DePue, J., 2004. Multiple behavioral risk factor interventions in primary care: summary of research evidence. Am. J. Prev. Med. 27 (2), 61–79. http://dx.doi.org/10.1016/j.amepre.2004.04.023.

Hale, D.R., Fitzgerald-Yau, N., Viner, R.M., 2014. A systematic review of effective interventions for reducing multiple health risk behaviors in adolescence. Am. J. Public Health 104 (5), e19–e41. http://dx.doi.org/10.2105/AJPH.2014.301874.

Harrington, J.M., Dahly, D.L., Fitzgerald, A.P., Gilthorpe, M.S., Perry, I.J., 2014. Capturing changes in dietary patterns among older adults: a latent class analysis of an ageing Irish cohort. Public Health Nutr. 17 (12), 2674–2686. http://dx.doi.org/10.1017/s1368980014000111.

Heroux, M., Janssen, I., Lee, D.C., Sui, X.M., Hebert, J.R., Blair, S.N., 2012. Clustering of unhealthy behaviors in the Aerobics Center Longitudinal Study. Prev. Sci. 13 (2), 183–195. http://dx.doi.org/10.1007/s11121-011-0255-0.

Hofstetter, H., Dusseldorp, E., van Empelen, P., Paulussen, T., 2014. A primer on the use of cluster analysis or factor analysis to assess co-occurrence of risk behaviors. Prev. Med. 67, 141–146. http://dx.doi.org/10.1016/j.ypmed.2014.07.007.

HSCIC, 2014. Statistics on Smoking, England 2014. HSCIC ([cited 2015 20 July]. Available from: http://www.hscic.gov.uk/catalogue/PUB14988).

Hyman, D.J., Pavlik, V.N., Taylor, W.C., Goodrick, G.K., Moye, L., 2007. Simultaneous vs sequential counseling for multiple behavior change. Arch. Intern. Med. 167 (11), 1152. http://dx.doi.org/10.1001/archinte.167.11.1152.

Keyes, K.M., Li, G., Hasin, D.S., 2011. Birth cohort effects and gender differences in alcohol epidemiology: a review and synthesis. Alcohol. Clin. Exp. Res. 35 (12), 2101–2112. http://dx.doi.org/10.1111/j.1530-0277.2011.01562.x.

King, K., Meader, N., Wright, K., Graham, H., Power, C., Petticrew, M., et al., 2015. Characteristics of interventions targeting multiple lifestyle risk behaviours in adult populations: a systematic scoping review. PLoS One 10 (1), e0117015. http://dx.doi.org/10.1371/journal.pone.0117015.

Kvaavik, E., Batty, G., Ursin, G., Huxley, R., Gale, C.R., 2010. Influence of individual and combined health behaviors on total and cause-specific mortality in men and women: the United Kingdom health and lifestyle survey. Arch. Intern. Med. 170 (8), 711–718. http://dx.doi.org/10.1001/archinternmed.2010.76.

Laaksonen, M., Prättälä, R., Karisto, A., 2001. Patterns of unhealthy behaviour in Finland. Cent. Eur. J. Public Health 11 (3), 294–300. http://dx.doi.org/10.1093/eurpub/11.3.294.

Leventhal, A.M., Huh, J., Dunton, G.F., 2014. Clustering of modifiable biobehavioral risk factors for chronic disease in US adults: a latent class analysis. Perspect. Public Health 134 (6), 331–338. http://dx.doi.org/10.1177/1757913913495780.

Linardakis, M., Smpokos, E., Papadaki, A., Komninos, I.D., Tzanakis, N., Philalithis, A., 2013. Prevalence of multiple behavioral risk factors for chronic diseases in adults aged 50+, from eleven European countries — the SHARE study (2004). Prev. Med. 57 (3), 168–172. http://dx.doi.org/10.1016/j.ypmed.2013.05.008.

Loef, M., Walach, H., 2012. The combined effects of healthy lifestyle behaviors on all cause mortality: a systematic review and meta-analysis. Prev. Med. 55 (3), 163–170. http://dx.doi.org/10.1016/j.ypmed.2012.06.017.

McAloney, K., Graham, H., Law, C., Platt, L., 2013. A scoping review of statistical approaches to the analysis of multiple health-related behaviours. Prev. Med. 56 (6), 365–371. http://dx.doi.org/10.1016/j.ypmed.2013.03.002.

Meng, Y., Holmes, J., Hill-McManus, D., Brennan, A., Meier, P.S., 2014. Trend analysis and modelling of gender-specific age, period and birth cohort effects on alcohol abstention and consumption level for drinkers in Great Britain using the General Lifestyle Survey 1984–2009. Addiction 109 (2), 206–215. http://dx.doi.org/10.1111/add.12330.

Mulder, M., Ranchor, A.V., Sanderman, R., Bouma, J., van den Heuvel, W.J., 1998. The stability of lifestyle behaviour. Int. J. Epidemiol. 27 (2), 199–207. http://dx.doi.org/10.1093/ije/27.2.199.

Muthen, M., 2014. Mplus: Version 7. Muthen & Muthen, Los Angeles, CA.

Nigg, C.R., Allegrante, J.P., Ory, M., 2002. Theory-comparison and multiple-behavior research: common themes advancing health behavior research. Health Educ. Res. 17 (5), 670–679. http://dx.doi.org/10.1093/her/17.5.670.

Nylund, K.L., Asparouhov, T., Muthén, B.O., 2007. Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo Simulation Study. Struct. Equ. Model. 14 (4), 535–569. http://dx.doi.org/10.1080/10705510701575396.

Paavola, M., Vartiainen, E., Haukkala, A., 2004. Smoking, alcohol use, and physical activity: a 13-year longitudinal study ranging from adolescence into adulthood. J. Adolesc. Health 35 (3), 238–244. http://dx.doi.org/10.1016/j.jadohealth.2003.12.004.

Parsons, T.J., Power, C., Manor, O., 2006. Longitudinal physical activity and diet patterns in the 1958 British Birth Cohort. Med. Sci. Sports Exerc. 38 (3), 547–554. http://dx.doi.org/10.1249/01.mss.0000188446.65651.67.

Poortinga, W., 2007. The prevalence and clustering of four major lifestyle risk factors in an English adult population. Prev. Med. 44 (2), 124–128. http://dx.doi.org/10.1016/j.ypmed.2006.10.006.

Power, C., Elliott, J., 2006. Cohort profile: 1958 British birth cohort (National Child Development Study). Int. J. Epidemiol. 35 (1), 34–41. http://dx.doi.org/10.1093/ije/dyi183.

Prochaska, J.J., Velicer, W.F., Nigg, C.R., Prochaska, J.O., 2008. Methods of quantifying change in multiple risk factor interventions. Prev. Med. 46 (3), 260–265. http://dx.doi.org/10.1016/j.ypmed.2007.07.035.

RCP, 2012. Fifty Years Since Smoking and Health. Royal College of Physicians, London (Available from: https://www.rcplondon.ac.uk/sites/default/files/fifty-years-smoking-health.pdf).

Room, R., 2004. Smoking and drinking as complementary behaviours. Biomed. Pharmacother. 58 (2), 111–115.

Sayon-Orea, C., Martinez-Gonzalez, M.A., Bes-Rastrollo, M., 2011. Alcohol consumption and body weight: a systematic review. Nutr. Rev. 69 (8), 419–431. http://dx.doi.org/10.1111/j.1753-4887.2011.00403.x.

Schneider, S., Huy, C., Schuessler, M., Diehl, K., Schwarz, S., 2009. Optimising lifestyle interventions: identification of health behaviour patterns by cluster analysis in a German 50 survey. Cent. Eur. J. Public Health 19 (3), 271–277. http://dx.doi.org/10.1093/eurpub/ckn144.

Schnuerer, I., Baumann, S., Haberecht, K., Gaertner, B., John, U., Freyer-Adam, J., 2015. Patterns of health risk behaviors among job-seekers: a latent class analysis. Int. J. Public Health 60 (1), 111–119. http://dx.doi.org/10.1007/s00038-014-0623-1.

Schoon, I., Parsons, S., 2003. Lifestyle and Health-Related Behaviour. Changing Britain, Changing Lives. Institute of Education Press, London, pp. 237–260.

Schuit, A.J., van Loon, A.J.M., Tijhuis, M., Ocké, M.C., 2002. Clustering of lifestyle risk factors in a general adult population. Prev. Med. 35 (3), 219–224. http://dx.doi.org/10.1006/pmed.2002.1064.

Shepherd, P., 2012a. 1958 National Child Development Study Ethical Review and Consent. Centre for Longitudinal Studies, London (Available from: http://www.cls.ioe.ac.uk/Publications.aspx?sitesectionid=93&).

Shepherd, P., 2012b. 1970 British Cohort Study Ethical Review and Consent. Centre for Longitudinal Studies, London (Available from: http://www.cls.ioe.ac.uk/page.aspx?&sitesectionid=1342&sitesectiontitle=Data+Linkage).

Silva, D.A.S., Peres, K.G., Boing, A.F., González-Chica, D.A., Peres, M.A., 2013. Clustering of risk behaviors for chronic noncommunicable diseases: a population-based study in southern Brazil. Prev. Med. 56 (1), 20–24. http://dx.doi.org/10.1016/j.ypmed.2012.10.022.

Singh, G.M., Micha, R., Khatibzadeh, S., Shi, P., Lim, S., Andrews, K.G., et al., 2015. Global, regional, and national consumption of sugar-sweetened beverages, fruit juices, and milk: a systematic assessment of beverage intake in 187 countries. PLoS One 10 (8), e0124845. http://dx.doi.org/10.1371/journal.pone.0124845.

Stead, M., Munafo, M., Fuller, E., Bauld, L., Angus, K., MacDonald, L., et al., 2013. Scoping and Feasibility Study to Develop and Apply a Methodology for Retrospective Adjustment of Alcohol Consumption Data. University of Stirling: Public Health Research Consortium (Available from: http://phrc.lshtm.ac.uk/papers/FINAL_REPORT_PHRC_Scoping_and_feasibility_study_-_alcohol_consumption_data_(October_2013).pdf).

Steptoe A. Evaluating strategies for behavior change to reduce cardiovascular risk. Nat. Clin. Pract. Cardiovasc. Med. 2007;4(11):598–9. http://dx.doi.org/10.1038/ncpcardio1003.

Sullivan, A., Brown, M., Sullivan, A., Brown, M., 2013. Overweight and Obesity in Mid-life: Evidence From the 1970 Birth Cohort Study at Age 42. Centre for Longitudinal Studies, London (Available from: http://www.cls.ioe.ac.uk/page.aspx?sitesectionid=795sitesectiontitle=Welcome+to+the+1970+British+Cohort+Study+(BCS70)).

Telama, R., 2009. Tracking of physical activity from childhood to adulthood: a review. Obes. Facts 2 (3), 187–195. http://dx.doi.org/10.1159/000222244.

Tobias, M., Jackson, G., Yeh, L.-C., Huang, K., 2007. Do healthy and unhealthy behaviours cluster in New Zealand? Aust. N. Z. J. Public Health 31 (2), 155–163. http://dx.doi.org/10.1111/j.1753-6405.2007.00034.x.

van Nieuwenhuijzen, M., Junger, M., Velderman, M.K., Wiefferink, K.H., Paulussen, T.W.G.M., Hox, J., et al., 2009. Clustering of health-compromising behavior and delinquency in adolescents and adults in the Dutch population. Prev. Med. 48 (6), 572–578. http://dx.doi.org/10.1016/j.ypmed.2009.04.008.

Verger, P., Lions, C., Ventelou, B., 2009. Is depression associated with health risk-related behaviour clusters in adults? Cent. Eur. J. Public Health 19 (6), 618–624. http://dx.doi.org/10.1093/eurpub/ckp057.

Vermeulen-Smit, E., Ten Have, M., Van Laar, M., De Graaf, R., 2015. Clustering of health risk behaviours and the relationship with mental disorders. J. Affect. Disord. 171, 111–119. http://dx.doi.org/10.1016/j.jad.2014.09.031.

Wang, J., Wang, X., 2012. Structural Equation Modeling: Applications Using Mplus. John Wiley & Sons.

Watts, P., Buck, D., Netuveli, G., Renton, A., 2015. Clustering of lifestyle risk behaviours among residents of forty deprived neighbourhoods in London: lessons for targeting public health interventions. J. Public Health http://dx.doi.org/10.1093/pubmed/fdv028.

WHO, 2014. Global Status Report on Non-communicable Diseases 2014. WHO, Geneva (Available from: http://www.who.int/global-coordination-mechanism/publications/global-status-report-ncds-2014-eng.pdf).

WHO, 2015. Guideline: Sugars Intake for Adults and Children. WHO, Geneva (Available from: http://www.who.int/nutrition/publications/guidelines/sugars_intake/en/).

Zacny, J.P., 1989. Behavioral aspects of alcohol-tobacco interactions. Recent developments in alcoholism: an official publication of the American Medical Society on alcoholism, the Research Society on alcoholism, and the National Council on. Alcoholism 8, 205–219.