Padgett, RN, Andretta, JR, Cole, JC, Percy, A, Sumnall, H and McKay, MT

Intervention impact on alcohol use, alcohol harms, and a combination of both: A latent class, secondary analysis of results from a randomized controlled trial

http://researchonline.ljmu.ac.uk/id/eprint/15688/

Article

Citation (please note it is advisable to refer to the publisher’s version if you intend to cite from this work)

Padgett, RN, Andretta, JR, Cole, JC, Percy, A, Sumnall, H and McKay, MT (2021) Intervention impact on alcohol use, alcohol harms, and a combination of both: A latent class, secondary analysis of results from a randomized controlled trial. Drug and Alcohol Dependence, 227. ISSN 0376-
Intervention impact on alcohol use, alcohol harms, and a combination of both: A latent class, secondary analysis of results from a randomized controlled trial.

Abstract

Background. Alcohol use among adolescents, as well as its associated harm, places a substantial burden on health, and public services more generally. To date, attempts to intervene at a universal level have yielded results varying from iatrogenic, to null, while in a few cases, skill-enhancing universal interventions have successfully impacted drinking behaviors. One such intervention is (*Intervention A*). The present study is a secondary data analysis from (*RCT/Int A*), providing new, and more nuanced findings. Methods. A total of 13,914 youth (41.7% female) participated in this cRCT where schools were randomly assigned to a control or intervention group. Growth mixture modelling was used to identify trajectory classes from baseline through third follow-up (+33 months) of youth on heavy episodic drinking (HED) and alcohol related health harms (ARH). Extracted classes were related to school intervention participation using multinomial logistic regression. Results. Five trajectory classes of the HED and ARH composite were identified: Low (62%), Late Onset (16%), Early Onset (13%), Delayed Onset (7%), and Unstable (3%). The intervention was most strongly related to Late Onset (OR=0.50, 95%CI [0.25, 1.01]) and Delayed Onset (OR=0.55, 95%CI [0.26, 1.16]), although not statistically significant. With classes constructed with ARH only, the Delayed Onset class was significantly related to the intervention (OR=0.60, 95%CI [0.43, 0.84]). Conclusions. These results support those previously reported on the *RCT/IntA* trial, and provide a more nuanced insight into the effects of the intervention.

Key words: Adolescent; Latent Class Analysis; Growth mixture modelling; Alcohol.
**Intervention impact on alcohol use, alcohol harms, and a combination of both: A latent class, secondary analysis of results from a randomized controlled trial.**

1. **Introduction**

   Alcohol use is a major contributor to the global disease burden and a risk factor for injuries and mortality, and this disproportionately affects young people (Shield et al., 2020). The overall proportion of adolescents drinking alcohol in the United Kingdom (UK) has declined in recent years, although alcohol-related health harms (ARH) remain high (Healey et al., 2014, Public Health England, 2019). School- and community-based prevention and education programmes are central to comprehensive responses for reducing ARH (Babor et al., 2010, Burton et al., 2017). There is some evidence, albeit limited, that school-based interventions can positively impact on drinking behaviour (mainly small-sized effects on frequency of consumption, amount consumed, and in particular effectiveness of individual brief interventions on consumption; Das et al., 2016; Strøm et al., 2014). Effective interventions include those that aim to develop social skills rather than those which simply seek to enhance knowledge (Faggiano et al., 2008), or that combine student- and parent-based activities (Newton et al., 2017).

   However, most community-based prevention interventions have not been subject to evaluation (Faggiano et al., 2014), and some may even produce iatrogenic effects (e.g. Allara et al., 2019).

   School-based prevention programmes are typically universal interventions that are delivered to populations regardless of levels of alcohol-related risk (Foxcroft, 2014). However, within populations there is variability in the extent and determinants of alcohol use. Epidemiological studies suggest that whilst there is some evidence of collectivity in the recent decline in youth drinking, the magnitude is determined by factors such as level of consumption,
age, gender, and socioeconomic status (Pape et al., 2018, Oldham et al., 2020). Examining the effects of prevention interventions, there is also evidence of sub-group effects, with universal interventions having differential impact on the basis of pre-intervention/baseline levels of risk factors, including alcohol use, and target audience(s) (for example, *Authors secondary analysis* reported counter-intuitive findings that the greatest impact of a classroom intervention was on baseline unsupervised drinkers, $M_{age} = 12.5$ years. Further, Koning et al., 2009 reported a small but significant reduction in initiation to heavy weekly drinking when both parents and children were engaged in an intervention, as opposed to children only). Given that those who report early initiation of alcohol intoxication are more likely to report adverse alcohol-related outcomes in young adulthood (Kuntsche et al., 2013; Maimaris and McCambridge, 2014; Morean et al., 2014), it is advantageous to understand whether universal interventions are effective across groups reporting different levels of alcohol-related risk.

The *RCT/Int A* trial (*Authors*) was a large (N $\approx$ 12,000 participants) Cluster Randomised Controlled Trial (cRCT) comparing the effects of a combined school- and parental-based intervention, based on the *School/Parenting interventions*. The version of *Intervention A* used in the *RCT/Int A* study had been adapted and piloted tested previously in a United Kingdom (UK) context (*Authors*).

The *RCT/INT A* trial included 70 post-primary schools in *Country A* and a further 35 post-primary schools in *Country B*. In the *RCT/Int A* trial, questionnaires were administered to participants at baseline (T0) and at three follow-ups: +12 months, +24 months, and primary outcome analyses were performed on data gathered at +33 months from baseline. The intervention was delivered between T1 and T2 (six lessons), and a further four lessons were
delivered between T2 and T3. This meant that data gathered to assess the primary outcomes was at least +10 months after full intervention delivery. The intervention group was compared to a control group who had continued to receive Education as Normal (EAN). Those in receipt of EAN reported significantly higher rates of Heavy Episodic Drinking (HED) in the past month (Primary outcome #1) than pupils in the intervention schools. However, the study arms did not differ significantly in terms of the number of ARH reported in the previous six months (Primary Outcome #2). The HED finding differed from the results of a pilot trial of *Intervention A*, where self-reported number of ARH were significantly reduced. However, in the pilot, the school component was only delivered in *Country A*, and to pupils who were one year older. In the present study, an extended measure (by one item) of ARH was employed.

While the primary analysis within the main cRCT did not find a significant overall intervention effect for ARH, exploratory subgroup analyses indicated a significant intervention effect on ARH when the sample was restricted only to those who reported early onset drinking (i.e. restricted to only those pupils who reported “ever” drinking at baseline T0, or only those who reported “last year” drinking at T0; *Authors*). This suggests that the intervention may have a differential effect on pupils who have different consumption patterns. This paper is an extension of this exploratory analysis. Rather than assessing alcohol use at T3 (+33 months) on the basis of a single measure at T0 (baseline), this study examines the effect of the intervention on the development of drinking patterns (trajectories) across all survey sweeps, that is from T0 to T3. It was hypothesised that consumption patterns characterised by lower, declining, or delayed HED or ARH would be more common amongst pupils in the Intervention schools than pupils in the Control schools.
2. Material and Methods

2.1 Participants

Participants were pupils in N = 105 High schools (N = 13,194; [N = 5499 females, 41.7%]) in *Country A* and *Country B*. Schools were recruited to participate in a cRCT, and were allocated to Intervention or Control arms. In addition to school children, the parents of children in schools randomized into the Intervention arm were targeted with a brief intervention. These interventions are detailed elsewhere (*Authors*) however, some key information is given below.

2.2 Intervention

The intervention was a classroom-based alcohol education intervention, coupled with a brief alcohol intervention for parents/carers. It combines a harm reduction philosophy with skills training, education and activities designed to encourage positive behavioural change. In addition to information and alcohol-related facts, the intervention is interactive both in terms of tasks (for example pouring units of alcohol) and discussion/debate. Briefly, the intervention consisted of ten classroom-based lessons delivered by trained schoolteachers over a two year period, with six lessons delivered in year one, and four lessons in year two. Lessons delivered in year one focused on a broad range of alcohol-related issues including, but not limited to: myths about alcohol; alcohol and the media; alcohol and the body; units of alcohol; the relationship between increasing levels of consumption and likely behavioural outcomes; as well as a look at some scenario-based situations. The lessons in year two focused on drinking contexts, and possible harms that might emerge from drinking in such contexts. Students were asked to focus on a particular ‘night out’ with a view to identifying possible harms and developing ameliorating
strategies. Additionally, students were encouraged to debate deliberately provocative statements, for example, ‘drinking vodka is worse than drinking beer’. As part of pre-intervention training events, teachers were encouraged to facilitate discussion where possible, rather than focusing on completion of the workbooks which accompanied the lessons. Materials were also provided in digital format in order to facilitate interactive delivery of the programme. In addition, the parents of the children in the Intervention group were invited to attend a single brief intervention, facilitated by an external organisation, in the school setting. This event focused on rule setting, and culminated in an agreed set of alcohol-specific rules to be applied across the homes of all those present. Only 9% and 2.5% of eligible parents attended the in-person events in Northern Ireland and Scotland respectively. Intervention children’s parents received a mailed reminder of the content of this session, regardless of whether they attended in person. A total of 31% and 18% of parents responded to a request therein to complete a questionnaire on parental rule-setting, in Northern Ireland and Scotland respectively.

2.3 Measures

Primary outcomes

This study is a secondary analysis of the two primary outcomes from *RCT/Int A*;

(i) The number of self-reported HED episodes in the previous 30 days (HED; defined as the consumption of ≥6 units of alcohol [males]/ ≥4.5 units [females]). HED was reported as 0, 1,…, 6 or more episodes.

(ii) The number of self-reported ARH (caused by own drinking) in the previous six months. Pupils were asked about the frequency of experiencing 16 types of harm (e.g. being sick after drinking, getting into trouble with your parents as a result of your drinking) in the last six
months. Responses for each harm were dichotomised (none/1 or more) and then summed to produce an overall count of the number of different harms experienced. Harms are listed in Supplementary Materials, part A.

To assess the HED primary outcome, participants were presented with pictorial prompts of how much alcohol ≥6/≥4.5 UK units represents. Pictures presented the most popular drinks consumed in the two study areas, and respondents were asked to report the frequency of consuming this amount of alcohol over the previous month. ARH were measured using a 16-item scale developed for the *RCT/Int A* trial (internal consistency 0.9; *Authors*). For example, participants were asked to report frequency of having a hangover after drinking, or if they had got into a physical fight when drinking.

Data were also collected on gender of the school (mixed/boys only/girls only), country (Northern Ireland/Scotland), and level of free school meals (FSM) entitlement within the school (a tertile split). FSM entitlement is often used as a proxy for socio-economic status (SES; Hobbs & Vignoles, 2010) within the UK.

2.4 Statistical Analyses

Missing Data. Missing data were investigated using R (R Core Team, 2019). We identified cases with missing data on the outcomes at all time points or were only present during one time point. These cases were omitted from further analysis due to not contributing information on the growth trajectory† (for notes please see Supplementary materials, part B). For all other cases, we imputed missing data using imputation by chained equations using the mice package (van Buuren and Groothuis-Oudshoorn, 2011). We used the covariates described above,
the HED and ARH measures, and a missing data indicator variable for each time point as variables that inform each chain about missingness.

Longitudinal Model. We used growth mixture modelling (GMM) to model longitudinal change in HED and ARH accounting for unobserved heterogeneity (latent classes) using Mplus version 8.3 (Muthén and Muthén, 2017). Models were estimated using maximum likelihood with robust standard errors with numerical integration. Using this approach, classes of individuals were identified based on similar growth patterns. Latent classes were constructed for HED and ARH simultaneously.

We initially proposed a parallel process model where a growth mixture models for HED and ARH were estimated simultaneously with the growth models sharing the latent class variable. However, two issues arose when fitting these models to our data. First, when accounting for the discrete nature of these data (i.e., as Poisson or negative binomial distributed count) this model required six dimensions of integration (one for each continuous latent variable) and was computationally intractable given our computing resources. Second, when assuming the outcomes where normally distributed, therefore not accounting for the discreteness, nearly all converged solutions contained negative variances or uninterpretable results.

Because of the issues of convergence and interpretability outlined above, we modeled HED and ARH using an independent and composite score approach. First, we model HED and ARH independently by fitting GMM to each outcome without the other as part of the model. Then, to capture part of the joint distribution of HED and ARH, we modeled a composite (sum) score of the two outcomes to identify growth trajectories. The model estimated followed the
structure as shown in Figure 1 that include repeated measurements with a latent growth model, a categorical latent variable, and covariates.

The number of classes (growth trajectories) was determined using several indices. We fitted one to six class solutions. We used two information criteria and two likelihood ratio tests to determine statistical measures of fit and to compare solutions. The two information criteria were used: (a) the Bayesian information criteria (BIC) and (b) the sample-size adjusted Bayesian information criteria (ssBIC); lower BIC and ssBIC values indicate a better fitting model (Nylund et al., 2007). The two likelihood ratio tests used were: (a) the Vuong-Lo-Mendell- Rubin (VLMR) LRT, and (b) the Lo-Mendell- Rubin (LMR) LRT; where the fit of the lower class solution was compared to a higher class solution to determine if the additional complexity adds to the fit of the model to these data. The final number of classes was selected based on the BIC, ssBIC, VLMR and LMR p-values, appropriateness of class sizes, parsimony, research questions, and substantive interpretability (Bauer & Curran, 2003).

Inference was focused on the individual level of growth patterns, meaning the clustering of children in schools was not of substantive interest. Therefore, we treated our analysis as a two-level model where the individual observations were nested within children. The clustering of children within schools was controlled for by maximum likelihood estimation with cluster robust standard errors (Muthén and Muthén, 2017). Responses were modeled as Poisson distributed counts, which accounted for the discrete nature of the observed data.

Figure 1 – about here

Figure 1. Path diagram of growth mixture model for HED and ARH composite.
Note. Covariate effects estimated with the three-step ML procedure. Error terms omitted for simplicity.

The resulting trajectory classes from the GMM were used to investigate how the intervention and other covariates related to change over time. The strength of the relationship between latent class and covariates was characterised using multinomial logistic regression analysis, assessing the likelihood of membership to one trajectory class compared with another. These analyses follow the maximum likelihood three-step approach to relation latent class membership to covariates (Nylund-Gibson et al., 2019).

3. Results

3.1 Descriptive Data

The descriptive data for the observed outcomes at each time point in each group are reported in Table 1. There was an increase in the average for each outcome in both the Intervention and Control groups and the standard deviation (variability) also increased over time. The increase in variability indicates that the similarity of participants within the intervention and control groups changed over time.

Table 1

3.2 Growth Mixture Model of HED

A four-class solution with quadratic growth was identified as optimal due to the low separation among classes at higher class solutions and negative factor variances which lead to difficulty in interpretation of effects (see supplement material for more information on class selection). Class 1 (6% of sample) was associated with highly variable use of alcohol over time.
(termed Unstable); Class 2 (79%), the largest group, with minor to no episodes of HED (Low Use); Class 3 (7%) with delayed onset increases in use (Delayed Onset); and Class 4 (8%) with variable but early onset in use (Early Onset). Using multinomial logistic regression, we examined the relationship between Intervention status and modal class membership, whilst controlling for demographic characteristics (sex, FSM, and country) (Table S-4). Results demonstrated a significant positive effect of Intervention on class membership so that participants receiving the Intervention were more likely to be in the Unstable than Early Onset class; Low Use than Delayed Onset or Early Onset classes; and Delayed Onset class rather than Early Onset class. Other comparisons are displayed in the Table S-4.

3.3 Growth Mixture Model of ARH

A five-class solution with quadratic growth was identified as optimal due to higher class solutions being uninterpretable and negative factor variances (see supplement material for more information on class selection). Class 1 (10%) was associated with Delayed Onsets in ARH over time (Delayed Onset); Class 2 (8%) with variable but linear increase in ARH (Early Onset); Class 3 (3%) with decreasing ARH (Declining ARH); Class 4 (67%) with low to no ARH (Low ARH); and Class 5 (12%) with Late Onsets in ARH (Late Onset). Multinomial logistic regression showed only two significant findings, namely, that participation in the Intervention was significantly associated with membership of the Delayed Onset, and Early Onset classes, compared to the Low class (Table S-8). There was greater strength of positive association between the Intervention condition and Delayed Onset class membership, suggesting that the Intervention may have had time-limited effects on ARH. Additionally, country was strongly related to membership in the Early Onset class. Upon inspection of the breakdown of modal class
by country, we found that blinded youth were disproportionately represented in the Early Onset class. This dominance of blinded youth in the Early Onset class may be indicative of a cultural difference in ARH between the two countries.

3.4 Growth Mixture Model of both HED and ARH

The five-class quadratic solution was identified as optimal (see Table S-10). Although the best statistical fit was found with the six-class solution, upon examining the trajectories in more detail we found that in the six-class solution, two of the classes were nearly identical except for a substantively unimportant difference in the linear slope. The five classes are described in Table 2 with the class specific parameters and characteristics. Class 1 was associated with delayed onset increases in HED and ARH over time (Delayed Onset); Class 2 late onset increases in HED and ARH (Late Onset); Class 3 increases approximately linearly early on but with high within class variation in HED and ARH (Early Onset); Class 4 low to no HED and ARH (Low); and Class 5 with highly variable use of alcohol over time (termed Unstable); The individual and class average growth patterns are shown in Figure 2.

Table 2 and Figure 2 – about here

Figure 2. Five class solution of composite of HED and ARH. The left panel shows the individual estimated growth curves over all participants in the five classes. The right panel displays the class average trajectories (i.e., based on the model parameters).

We found that most Intervention participants were identified as belonging to the Low Class; however, this is due to the large size of this class (62% of the sample; see Table 6). Using

12
multinomial logistic regression controlling for Sex, FSM and Country, we found that participation in the Intervention was not statistically associated with class membership when comparing odds between classes (Table 3).

4. Discussion

We had hypothesized that consumption patterns characterised by lower, declining, or delayed HED or ARH would be more common amongst pupils in the Intervention schools than pupils in the Control schools, and while there was some evidence that this was the case, it was notable that results were either small-sized in nature, or not statistically significant.

At a broad level, the results of the present study for HED are broadly in line with those observed in (*Authors blinded 1 & 2*), so that participation in *Intervention A* is associated with a significant reduction in drinking behavior. Whereas results reported in the cRCT previously (RCT) showed no intervention effect for ARH, the present analyses suggested two small effects. Compared to those in the Low class, those in the Delayed Onset and Early Onset classes were significantly more likely to have been in the Control group. These findings require further investigation, but as data were collected over four waves, with the intervention being delivered in two phases, between waves 1 and 2, and between waves 2 and 3, there is a possibility that the results for the Early Onset class represents an early spike in Control group drinking behaviors which was protected against by intervention effects. The fact that Intervention exposure was associated with Low class membership, compared to Delayed Onset suggests that effects were longer lasting than the intervention period itself. This was also observed in a separate analysis using participants who participated in two further data sweeps (*Authors*), where the Intervention and control groups differed at +33 months on HED, but
there was no significant effect of Intervention 30 months after Intervention cessation. This is consistent with other prevention trials that have incorporated long-term follow up, and most effective interventions only have relatively short-term effects (e.g. see reviews by (Champion et al., 2013, Foxcroft and Tsertsvadze, 2012; Newton et al., 2017)).

The lack of a significant finding for the classes derived from the model of the composite of HED and ARH was disappointing. We had previously hypothesized (*authors, blinded*) that the effect might be delayed, as a positive Intervention effect on these outcomes had been observed in a one-year older cohort in our pilot study. However, this hypothesis was not subsequently supported in a follow up study (*authors blinded*), and combined with the findings of the present analysis, suggests that a single delivery of *Intervention A* has no long-term effect on ARH. Despite holding intuitive appeal, there has been relatively little investigation of the impact of the inclusion of ‘booster’ sessions that aim to refresh the main content of the intervention and enhance retention and recollection of key concepts and/or continue with further age appropriate material as the young people mature and their needs develop. Furthermore, as well as *Intervention* being considered a complex intervention, it is also delivered into the complex environment of a school and community which have their own health and risk promoting characteristics and dynamics (Bonell et al., 2019, Bonell et al., 2013, Jamal et al., 2013). Any future implementations of *Intervention* should therefore consider the inclusion of booster sessions, and more detailed study of the school environment.

Other research on alcohol prevention interventions with young drinkers have also reported complex, equivocal, and counterintuitive findings. In one study examining the effects of responsible drinking messages, Moss et al (2017) reported that while passive exposure to
messages led to negative attitudes towards drunkenness, active exposure had the opposite effect. The authors also identified an increase in longer-term intentions to get drunk at age 18 when participants were passively exposed to responsible drinking messages. Research with older drinkers has also shown mixed results with iatrogenic effects in one study (Moss et al., 2015), and positive effects elsewhere (De Graaf et al., 2015; Glock et al., 2015). Moss et al (2017) suggested that factors such as active or passive engagement in health promotion interventions may have differential effects on different outcomes related to decisions around alcohol use. This, in addition to variability in interpersonal alcohol use experiences, and variability in programme implementation fidelity, might go some way to explaining the lack of effect for ARH, simultaneous to an effect for HED.

One further possibility that may explain the lack of effect for ARH, yet an effect for HED, is the inclusion of a parental component in (*RCT/Int A*), which had not been the case in the pilot study. Although studies have found that the effectiveness of school-based universal programs may be increased by the inclusion of parent-based components (Koning et al., 2013; Koutakis et al., 2008; Newton et al., 2017; Stormshak et al., 2011), there remains a lack of evidence about the indirect effects of interventions targeting parental alcohol use upon children’s use behaviours (Kuntsche and Kuntsche, 2016). It was not possible to disentangle the effect on student outcomes of parental participation in the intervention (approximately 75% of parents did not attend the parental component). Alcohol-specific parenting practices, including parents’ and the wider family’s own consumption, the communication of norms and expectations around alcohol within the (social) environment, and setting of authoritative rules around access and use, have all been shown to be an important predictors of childrens’ use (Cranford et al., 2010;
Handley and Chassin, 2013, Koning et al., 2012, Shaw et al., 2018). It may be that a lack of engagement of the majority of parents with the intervention may have been passively (or actively) communicated towards children, and this may have led to an under-reporting of ARH in this group.

5. Limitations

Three limitations should be noted. First, we were unable to estimate the parallel process model that jointly estimated the latent class with HED and ARH with separate growth models. We overcame this by creating a composite score. However, by weighting both indicators equally we may have masked unique growth trajectories that would have emerged if the full model was investigated. For example, a pattern of increasing in HED and decreasing in ARH as drinkers learnt to avoid the latter through experience may be present but we could not identify this with our methods. Secondly, we imputed missing data using the available data. We reran our models using all available cases (including cases with data at only one time point) without imputing data, relying on the fact that maximum likelihood with robust standard errors (MLR) averages over missing data and treats missing data as missing at random (Little and Rubin, 2002). When we reran the model, we did not identify any substantive changes in the trajectory classes. However, the missing data mechanism may not have been missing at random and imputation of missing data may not have accounted for a key feature of the missing data which could have influenced the identification of trajectory classes. Third, the intervention may not have been equally effective for all participants within derived classes. This would mean that the intervention’s relationship to the trajectory classes is likely underestimated given the non-compliers in the study. Future research should focus on whether intervention compliance would
lead to more substantial relationships between the intervention and the latent classes. Finally, these findings were subject limitations that are common to alcohol prevention research. Firstly, the study was implemented in two countries of the UK, and so we make no assumptions about generalisability to other geographies. Secondly, alcohol use and ARH were self-reported, and although studies have shown this method to be generally reliable, although under-reporting may occur in heavier use adolescents (Lintonen et al., 2004, Northcote and Livingston, 2011), and a recanting analysis (*authors blinded*) found that contradictory reports did not affect primary outcome findings, it remains that alcohol use was not independently verified. However, for a large trial of this type, it was impractical to include biomarkers to verify alcohol use.

A strength of the methods employed in this study is that we investigated the relationship between trajectory classes and Intervention participation. We found that being in the Intervention was not significantly related to trajectory class membership meaning that on average we had a proportionally representative number of Intervention participants in each class. Using this knowledge, we can move to investigating how the Intervention related to changes in the trajectory within classes. Because we found that the Intervention was not meaningfully related to membership on average, future investigations of changes within trajectory classes will not be confounded by between class Intervention membership representation.
References


Northcote, J., Livingston, M., 2011. Accuracy of Self-Reported Drinking: Observational Verification of ‘Last Occasion’ Drink Estimates of Young Adults. Alcohol Alcohol. 46, 709-713.


Table 1.

Descriptive statistics of sample *Values show mean (SD)*

<table>
<thead>
<tr>
<th>Time Invariant Characteristics</th>
<th>N(^1)</th>
<th>Percent(^2)</th>
<th>% Missing</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>13,194</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex (Female)</td>
<td>5499</td>
<td>41.7</td>
<td>15.7</td>
</tr>
<tr>
<td>Free School Meals (Eligible)</td>
<td>2548</td>
<td>19.3</td>
<td>26.0</td>
</tr>
<tr>
<td>Country (blinded)</td>
<td>5176</td>
<td>39.2</td>
<td>0.0</td>
</tr>
<tr>
<td>Study Arm (Intervention)</td>
<td>6570</td>
<td>49.8</td>
<td>0.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Group</th>
<th>Baseline (T0)</th>
<th>12 Months (T1)</th>
<th>24 Months (T2)</th>
<th>33 Months (T3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HED</td>
<td>Control</td>
<td>0.14 (0.62)</td>
<td>0.19 (0.73)</td>
<td>0.30 (0.91)</td>
<td>0.53 (1.16)</td>
</tr>
<tr>
<td></td>
<td>Intervention</td>
<td>0.14 (0.57)</td>
<td>0.14 (0.57)</td>
<td>0.21 (0.74)</td>
<td>0.39 (1.05)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>0.14 (0.59)</td>
<td>0.16 (0.65)</td>
<td>0.25 (0.83)</td>
<td>0.46 (1.11)</td>
</tr>
<tr>
<td>ARH</td>
<td>Control</td>
<td>0.18 (0.63)</td>
<td>0.25 (0.81)</td>
<td>0.45 (1.09)</td>
<td>0.74 (1.36)</td>
</tr>
<tr>
<td></td>
<td>Intervention</td>
<td>0.17 (0.63)</td>
<td>0.21 (0.70)</td>
<td>0.40 (1.06)</td>
<td>0.65 (1.33)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>0.18 (0.63)</td>
<td>0.23 (0.75)</td>
<td>0.42 (1.07)</td>
<td>0.69 (1.34)</td>
</tr>
<tr>
<td>HED + ARH</td>
<td>Control</td>
<td>0.31 (1.07)</td>
<td>0.43 (1.35)</td>
<td>0.73 (1.80)</td>
<td>1.24 (2.30)</td>
</tr>
<tr>
<td></td>
<td>Intervention</td>
<td>0.31 (1.05)</td>
<td>0.33 (1.09)</td>
<td>0.60 (1.60)</td>
<td>1.24 (2.16)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>0.31 (1.05)</td>
<td>0.38 (1.23)</td>
<td>0.67 (1.70)</td>
<td>1.13 (2.24)</td>
</tr>
</tbody>
</table>

*Note.* \(^1\)Sample size prior to excluding cases with missing data on the outcomes at all time points or were only present during one time point, the final retained sample size was \( N = 12,333 \)

\(^2\)Percent excludes missing values.
Table 2.

5-Class solution details of HED plus ARH model

<table>
<thead>
<tr>
<th></th>
<th>Class 1</th>
<th>Class 2</th>
<th>Class 3</th>
<th>Class 4</th>
<th>Class 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Delayed Onset</strong></td>
<td><strong>Late Onset</strong></td>
<td><strong>Early Onset</strong></td>
<td><strong>Low</strong></td>
<td><strong>Unstable</strong></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>833 (7%)</td>
<td>1936 (16%)</td>
<td>1565 (13%)</td>
<td>7586 (62%)</td>
<td>413 (3%)</td>
</tr>
<tr>
<td>Female</td>
<td>473 (57)</td>
<td>1039 (54)</td>
<td>799 (51)</td>
<td>3692 (49)</td>
<td>127 (31)</td>
</tr>
<tr>
<td>Eligible for FSM</td>
<td>288 (35)</td>
<td>534 (28)</td>
<td>616 (39)</td>
<td>1704 (23)</td>
<td>136 (33)</td>
</tr>
<tr>
<td>from blinded</td>
<td>494 (59)</td>
<td>805 (42)</td>
<td>971 (62)</td>
<td>2319 (31)</td>
<td>179 (43)</td>
</tr>
<tr>
<td>Intervention Group</td>
<td>401 (48)</td>
<td>870 (45)</td>
<td>705 (45)</td>
<td>3953 (52)</td>
<td>204 (49)</td>
</tr>
<tr>
<td>Control Group</td>
<td>432 (52)</td>
<td>1066 (55)</td>
<td>860 (55)</td>
<td>3633 (48)</td>
<td>209 (51)</td>
</tr>
</tbody>
</table>

Average ARH across Time **Mean (SD)**

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>12 Months</th>
<th>24 Months</th>
<th>33 Months</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
<td>0.0 (0.0)</td>
<td>0.0 (0.2)</td>
<td>1.5 (1.9)</td>
<td>0.0 (0.2)</td>
</tr>
<tr>
<td><strong>12 Months</strong></td>
<td>0.3 (0.5)</td>
<td>0.1 (0.2)</td>
<td>2.9 (2.3)</td>
<td>0.0 (0.1)</td>
</tr>
<tr>
<td><strong>24 Months</strong></td>
<td>4.0 (2.1)</td>
<td>0.2 (0.4)</td>
<td>3.4 (2.7)</td>
<td>0.0 (0.2)</td>
</tr>
<tr>
<td><strong>33 Months</strong></td>
<td>3.7 (2.9)</td>
<td>2.9 (2.1)</td>
<td>4.0 (3.2)</td>
<td>0.00 (0.0)</td>
</tr>
</tbody>
</table>

Growth Model Parameters **Estimate (SE)**

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>Linear Growth Factor</th>
<th>Quadratic Growth Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intercept</strong></td>
<td>-6.63 (0.08)</td>
<td>-3.01 (0.05)</td>
<td>0.27 (0.01)</td>
</tr>
<tr>
<td><strong>Linear Growth Factor</strong></td>
<td>6.63 (0.08)</td>
<td>-1.47 (0.06)</td>
<td>0.71 (0.01)</td>
</tr>
<tr>
<td><strong>Quadratic Growth Factor</strong></td>
<td>-1.45 (0.02)</td>
<td>1.03 (0.02)</td>
<td>-0.15 (0.01)</td>
</tr>
</tbody>
</table>

**Note.** N = number of cases in each class where total sample size (N) is 12,333. 1Latent variable are on log scale which are exponentially related to the observed variables. Values in parentheses are standard errors (SE).
Table 3.

Results from multinomial logistic regression for composite of HED and ARH model

<table>
<thead>
<tr>
<th>Effect</th>
<th>Delayed Onset vs. Late Onset</th>
<th>Delayed Onset vs. Early Onset</th>
<th>Delayed Onset vs. Low</th>
<th>Delayed Onset vs. Unstable</th>
<th>Late Onset vs. Early Onset</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>95% CI</td>
<td>95% CI</td>
<td>95% CI</td>
<td>95% CI</td>
<td>95% CI</td>
</tr>
<tr>
<td>Intervention†</td>
<td>1.10</td>
<td>0.56</td>
<td>2.17</td>
<td>1.02</td>
<td>0.57</td>
</tr>
<tr>
<td>Sex‡</td>
<td>1.53</td>
<td>0.97</td>
<td>2.41</td>
<td>1.91*</td>
<td>1.30</td>
</tr>
<tr>
<td>FSM††</td>
<td>1.58</td>
<td>0.99</td>
<td>2.50</td>
<td>0.82</td>
<td>0.53</td>
</tr>
<tr>
<td>Country‡‡</td>
<td>9.62*</td>
<td>5.47</td>
<td>16.92</td>
<td>0.99</td>
<td>0.60</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Effect</th>
<th>Late Onset vs. Low</th>
<th>Late Onset vs. Unstable</th>
<th>Early Onset vs. Low</th>
<th>Early Onset vs. Unstable</th>
<th>Low vs. Unstable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>95% CI</td>
<td>95% CI</td>
<td>95% CI</td>
<td>95% CI</td>
<td>95% CI</td>
</tr>
<tr>
<td>Intervention†</td>
<td>0.50</td>
<td>0.25</td>
<td>1.01</td>
<td>0.85</td>
<td>0.45</td>
</tr>
<tr>
<td>Sex‡</td>
<td>1.75*</td>
<td>1.07</td>
<td>2.86</td>
<td>4.67*</td>
<td>2.97</td>
</tr>
<tr>
<td>FSM††</td>
<td>1.33</td>
<td>0.88</td>
<td>2.02</td>
<td>0.63*</td>
<td>0.42</td>
</tr>
<tr>
<td>Country‡‡</td>
<td>7.70*</td>
<td>4.63</td>
<td>12.79</td>
<td>0.36*</td>
<td>0.22</td>
</tr>
</tbody>
</table>

*p < .05, CI = 95% Wald confidence interval, OR = odds ratio, where the odds ratio reflects the odds of being in the class first named compared to class named second. FSM, Free School Meals; †Reference category: Control; ‡Reference category: male; ††Reference category: not being in receipt of FSM; ‡‡Reference category: blinded.