ROYAL SOCIETY OPEN SCIENCE

The impact of the COVID-19 pandemic on adolescent mental health: a natural experiment

Journal:	*Royal Society Open Science*
Manuscript ID	RSOS-211114.R1
Article Type:	Registered Report - Stage 1
Date Submitted by the Author:	02-Aug-2021
Complete List of Authors:	Santos, Joao; University of Manchester Mansfield, Rosie; University College London Deighton, Jessica; Anna Freud National Centre for Children and Families Hayes, Daniel; Anna Freud National Centre for Children and Families; King's College London Velikonja, Tjasa; Anna Freud National Centre for Children and Families Boehnke, Jan; University of Dundee Patalay, Praveetha; University College London; University College London; University College London
Subject:	health and disease and epidemiology < BIOLOGY, psychology < BIOLOGY
Keywords:	COVID-19, adolescence, mental health, depression, externalising, wellbeing
Subject Category:	Science, Society and Policy

R. Soc. open sci. article template

ROYAL SOCIETY OPEN SCIENCE

R. Soc. open sci.
doi:10.1098/not yet assigned

The impact of the COVID-19 pandemic on adolescent mental health: a natural experiment

Joao Santos[+ 1], Rosie Mansfield[+ 2], Jessica Deighton[3], Daniel Hayes[3, 4], Tjasa Velikonja[3], Jan R. Boehnke[^ 5], Praveetha Patalay[^*† 2, 6]

[1] Manchester Institute of Education, University of Manchester, Manchester, UK
[2] Centre for Longitudinal Studies, Social Research Institute, University College London, London, UK
[3] Evidence Based Practice Unit, Anna Freud National Centre for Children, and Families, London, UK
[4] Health Service and Population Research, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK
[5] School of Health Sciences, University of Dundee, Dundee, UK
[6] MRC Unit for Lifelong Health and Ageing, Population Science and Experimental Medicine, University College London, London, UK
[+] joint first
[^] joint last

Keywords: COVID-19, adolescence, mental health, depression, externalising, wellbeing

1. Summary

Background: Despite widespread concern about the impact of COVID-19 on adolescent mental health, there remains limited empirical evidence which can causally attribute changes to the pandemic. The current study aimed to overcome existing methodological limitations by exploiting a serendipitously occurring natural experiment within two ongoing, multi-phase cluster randomised controlled trials.
Method: Depressive symptoms (primary outcome), externalising difficulties and life satisfaction (secondary outcomes) were assessed at baseline (phase 1 [pre-COVID-19 group]: September - October 2018, phase 2 [COVID-19 group]: September - October 2019) and 1-year follow up (pre-COVID-19 group: January - March 2020, COVID-19 group: February - April 2021). Participants in phase 1 (N = xxxx*) acted as controls. In phase 2, participants (N = xxxx*) were exposed to the COVID-19 pandemic between the baseline and follow up assessments providing a natural experimental design. The primary analysis will use a random intercept linear multivariable regression model with phase (exposure to the COVID-19 pandemic) included as the key predictor and individual and school-level variables as covariates.
Results:
Conclusion:
* Participant numbers to be included at Stage 2 submission when the dataset is available (see shell flow diagram Figure 2. for greater details on how these will be estimated).

* Corresponding author: Praveetha Patalay, p.patalay@ucl.ac.uk
†Address: Centre for Longitudinal Studies, Social Research Institute, University College London, London, UK
MRC Unit for Lifelong Health and Ageing, Population Science and Experimental Medicine, University College London, London, UK

2. Introduction

2.1. Background

Prior to the COVID-19 pandemic, there was widespread concern over rising mental health difficulties experienced by adolescents. In 2017, between 14-17% of adolescents aged 11-19 were found to meet diagnostic criteria for at least one mental disorder in England (1). Cross-cohort studies have demonstrated increases in internalising difficulties that indicate a deterioration of adolescent mental health over time (2, 3). It is important to understand whether the COVID-19 pandemic has contributed further to increased mental health difficulties in adolescence.

Despite widespread concern and media coverage about the impact of COVID-19 and related school closure on adolescent mental health (4), there remains limited robust empirical evidence which can causally attribute mental health changes to the pandemic (5, 6). To isolate the pandemic's effect, studies must include pre-pandemic assessments of symptoms (7) and account for age effects given known developmental patterns in mental health difficulties (8). Even when longitudinal data is available, results must be considered in the context of secular trends in child and adolescent mental health (9). Differentiating between age or developmental changes and the impact of the COVID-19 pandemic is of particular relevance for younger populations, as internalising symptoms are known to increase year-on-year from mid-adolescence (10, 11).

Much of the existing evidence is based on cross-sectional studies with no pre-pandemic assessments of mental health. Longitudinal data for this population is limited and in this age group, poses the unique challenge of differentiating between developmental change and COVID-19 impact. A living systematic literature review investigating the changes in mental health symptoms within the same individuals from pre-COVID-19 and across distinct phases of the pandemic, identified only four studies with child and adolescent samples (by June 2021 when this was written), none of which were from the United Kingdom (UK) (12). Findings from these few studies are mixed, with increased internalising symptoms reported in Australia (13) and increased conduct and overall difficulties reported in Spain (14). In contrast, fewer depressive and externalising symptoms were reported in China (15) and the Netherlands respectively (8). More recently, results from a longitudinal, population-based study in Iceland revealed trajectories of pre-pandemic depressive symptoms between 2016 and 2018 and during the COVID-19 pandemic (16). Adolescents aged 13-18 years reported significantly more depressive symptoms during the pandemic, and mental wellbeing decreased beyond what might be expected based on existing time trends of adolescent mental health (6).

In the UK, data from an ongoing regional cohort (Wirral Child Health and Development Study), revealed stark increases in young adolescents' depressive symptoms, post-traumatic stress disorder, and externalising difficulties during the COVID-19 pandemic (17). At a national level, the only population-based data indicating changes in mental health has come from the COVID-19 follow-up of the 2017 prevalence study (18). A higher proportion of children were found to be experiencing mental health difficulties, albeit methodological limitations around low response rates and differences in the mode and method of assessment before and during the pandemic (19). Both these UK studies are limited in their ability to separate developmental changes from pandemic related impact.

2.2. Objectives

To address some of these methodological challenges, this study exploited the serendipitous design of two large, ongoing multi-phase intervention trials. Using two cohorts of students, we are better able to isolate the impact of the COVID-19 pandemic on adolescent mental health from age and longer-term trends than previous studies (see Figure 1). In each study-"phase", baseline measures were assessed with adolescents aged 11-15 years across secondary schools. While in both phases the baseline assessment happened prior to the pandemic, the follow-up in phase 1 (pre-COVID-19 group) was assessed just before the pandemic (January - March 2020). Adolescents participating in phase 2 (COVID-19 group), experienced the pandemic and school closures between baseline and follow-up (February - April 2021). Hence, this paper aims to answer the following research question: What impact has the COVID-19 pandemic had on adolescent mental health, specifically, depressive symptoms (primary outcome), externalising difficulties, and life satisfaction (secondary outcomes)? We hypothesise that after controlling for baseline variables, levels in depressive symptoms and externalising difficulties will be higher, and life satisfaction lower, during the COVID-19 pandemic compared to before.

There has been some evidence from UK studies tracking families throughout the pandemic, that children with special educational needs (SEN) and from low-income homes were particularly impacted by

COVID-19 related school closures and lockdown (20). To investigate whether this is the case in our study population, we will subsequently examine whether there were socio-demographic differences (based on gender, ethnicity, socio-economic disadvantage, and special educational needs) in the impact of the COVID-19 pandemic on adolescent mental health outcomes.

<Insert Figure 1 here>

3. Materials and Methods

3.1. Study Design and Setting

Details of the trials from which these data are drawn are described below. The Education for Wellbeing Programme (EfW) is an evaluation of five school-based, mental health and wellbeing interventions which are organised into two parallel group cluster randomised controlled trials (RCTs) (21, 22). Interventions in Schools for Promoting Mental Wellbeing: Research in Education (INSPIRE), is a four-arm cluster RCT comparing three interventions (Mindfulness, Relaxation and Strategies for Safety and Wellbeing) to usual provision (control). Approaches for Wellbeing and Mental Health Literacy: Research in Education (AWARE) is a cluster RCT consisting of three-arms, comparing two mental health education interventions (Youth Aware of Mental Health (YAM) and The Guide) to usual provision (control). Randomisation of schools was conducted following baseline data collection by King's Clinical Trials Unit using an equal allocation ratio (1:1:1). Minimisation was applied for deprivation (free school meal eligibility), geographic region (London, Greater Manchester and North-West, Bath & Bristol and Durham), urban/rural location, and mental health provision reported at baseline (prior interventions coded absent/present) to ensure that conditions were comparable. For a full description of interventions, study design and measures, see the trial protocol papers (21, 22).

Due to the size of the two trials, schools were recruited in two phases (Figure 1; allocation to all interventions in both phases). Outcomes were assessed at baseline (prior to intervention randomisation) (phase 1 [pre-COVID-19 group]: September - October 2018, phase 2 [COVID-19 group]: September - October 2019) and 1-year follow up (9-12 months after interventions had been delivered) (pre-COVID-19 group: January – 11th March 2020, COVID-19 group: February - April 2021). Participants in phase 1 (N =xxxx from XX schools) acted as controls. Those in phase 2 (N = xxxx from XX schools) were exposed to the COVID-19 pandemic between the baseline and follow up assessments, leading to a natural experiment.

<Insert Figure 2 here>

3.2. Participants

Recruitment of participants was conducted in multiple stages. First, schools selected groups in relevant year groups to receive an intervention, if allocated. Second, letters were sent to the parents/carers of these pupils with information about the study; at this stage, they were offered the chance to opt their child out of the research. Finally, before completing the online surveys, pupils were presented with an information sheet and could assent to taking part by ticking all relevant boxes. If assent was not gained, the young person could not be part of the evaluation. The first young person participated on 17 September 2018. Ethics approval was obtained from University College London Research Ethics Committee [6735/009, 6735/014].

The main analytic sample in the current study is defined as all schools that were recruited to the trial and that took part in pupil surveys at both timepoints (baseline and 1-year follow up). All participants who completed some items of the survey at baseline or 1-year follow up were considered as part of the primary analysis sample. Figure 2 illustrates the participant flow diagram. We cannot complete all boxes at this stage and are including this figure to demonstrate how we will represent the participant allocation and drop out at various stages and the final analytic sample.

3.3. Variables

3.3.1. Individual level covariates

We will examine group differences and control for a range of individual pupil-level characteristics. These include age group (school year 7, 8 or 9 at baseline), child gender (male or female), socio-economic position assessed using eligibility for free school meals (FSM eligible or not), ethnicity (white or ethnic minority) and

special education needs status (SEN, yes or no). Year group and gender data were provided directly by the schools and all other information on pupil-level socio-demographic characteristics will be available via linkage to the National Pupil Database (NPD) (see Table 1). This linkage has not yet occurred, but we expect that covariate data will be missing for a small proportion of children due to unsuccessful linkage and these children will be excluded from analysis (as shown in Figure 2.).

3.3.2. School level covariates

When investigating the effect of condition (pre-COVID-19 group vs. COVID-19 group) on adolescent mental health outcomes, we will also control for several school-level characteristics that were used for minimisation for randomisation following baseline data collection. This information was obtained from the Department for Education's, Get Information About Schools (GIAS) service. School-level free school meal eligibility (%) will be included as an indicator of deprivation, as well as urban/rural status. The extent of existing mental health provision (prior interventions) reported by schools at baseline will also be included as a covariate. See Table 1 for the source and coding for each school-level variable.

3.4. Measures

3.4.1. Outcome measures

At all timepoints and across both phases of the study, schools were instructed to administer the pupil questionnaires via a secure online survey in teacher facilitated sessions during the normal school day. As described in the introduction, the primary outcome was depressive symptoms with externalising difficulties and life satisfaction considered secondary outcomes.

3.4.2. Primary outcome: depressive symptoms

The primary outcome measure of this study is adolescent self-reported depressive symptoms. Participants completed the Short Mood and Feelings Questionnaire (SMFQ) (23), a 13-item self-report measure of depressive symptoms in the previous two weeks. Examples of questions included are 'I felt miserable or unhappy' and 'I felt I was no good anymore'. Questions are rated on a 3-point Likert scale (0= "not true", 1= "sometimes", 2= "true"). Possible scores range from 0 to 26, with higher scores indicating greater depressive symptoms. The continuous depressive symptoms score will be the primary outcome measure. A binary score to assess impact on prevalence of caseness will be generated using the established cut off score (≥ 12) indicating high levels of depressive symptoms (23).

3.4.3. Secondary outcomes: externalising difficulties and life satisfaction

Externalising difficulties were measured using the behavioural difficulties subscale of the Me and My Feelings questionnaire (24, 25), a 6-item self-report scale (e.g., 'I hit out when I'm angry') with three response options: "never", "sometimes", and "always". Responses were summed to create a total behavioural difficulties score, with higher scores indicating greater difficulties. The scale has an established cut-off score of ≥ 6, which will be used for analysis examining a binary outcome of high externalising difficulties (24).

Life satisfaction was measured using the Huebner Life Satisfaction Scale (LSS; 26). The original scale consisted of 10 items with four response options: "never", "sometimes", "often", and "almost always". In this study, we use the adapted version of the scale that was reduced to 7 items with 6-point Likert scales ranging from "strongly disagree" to "strongly agree" following psychometric testing (26, 27). Items include 'my life is going well' and two items, 'I would like to change many things in my life' and 'I wish I had a different kind of life' that will be reverse scored so that high scores indicate greater life satisfaction. A total score is created summing responses from the 7 items, with higher total scores indicating greater life satisfaction.

<Insert Table 1 here>

3.5. Analysis Strategy

3.5.1. Descriptive statistics and data checks

We will compare the distribution of baseline characteristics for the schools across the two phases, using effect sizes to describe potentially relevant univariate differences. The same comparison will be performed for participants. Descriptive statistics (means for continuous and percentages for binary outcomes) with 95% confidence intervals will be presented for the primary and secondary outcomes at baseline and follow-up for each phase. We will also present histograms showing the distributions of the continuous outcome variables at baseline and follow-up for each phase in the supplementary material.

<See Table 3 for an example of a prevalence table displaying the balance of school-level characteristics across phase 1 (pre-COVID-19 group) and phase 2 (COVID-19 group) at baseline and 1-year follow-up. >

3.5.2. Estimating the impact of the pandemic

All analyses will be conducted using Stata 17 software. The primary outcome analysis will use a random intercept (for schools) linear multivariable regression model with depressive symptoms at 1-year follow-up as the dependent variable. The model will be specified as follows:

Level 1: $\boldsymbol{DS_{ij} = \beta_{0j} + \beta_{(1-k)} IndivIV_{ij} + e_{ij}}$
Level 2: $\boldsymbol{\beta_{0j} = \gamma_{00} + \beta_{01} Phase_j + \beta_{0(2-l)} SchoolIVs_j + u_{0j}}$

with $\boldsymbol{DS_{ij}}$ as the depressive symptom score at 1-year follow up of student *i* in school *j*. On Level 1, the depression score is regressed upon fixed effects of the *k* individual student covariates (*IndivIV*; see Table 1); and on Level 2 on *l* school characteristics (see Table 1) with school-level variance $\boldsymbol{u_{0j}}$ and individual-level error term $\boldsymbol{e_{ij}}$. As all Level-1 slopes are defined as fixed effects, the additional specification equations are omitted. *Phase* (exposure to COVID-19 pandemic) is the key regressor and the primary result will be the coefficient $\boldsymbol{\beta_{01}}$, which, if found statistically significant at $p < .05$, will be interpreted as the COVID-19 pandemic having potentially had an effect on adolescents' depressive symptoms. The direction of the potential effect is "increased adolescent depressive symptoms" if the SMFQ score is higher in phase 2 (and "decreased" if higher in phase 1). If the coefficient is not statistically significant, the primary outcome analysis will be interpreted as "no supporting evidence for a difference was found". For the dichotomous high depressive symptoms and externalising symptoms, we will run similar random intercept logistic multivariable regression models and report odds ratios.

The same strategy will be employed for the secondary outcomes, but we will be clear that the primary outcome analysis takes priority in interpretation of the results. As the selection of independent variables is determined by the variables available in the main study, the reporting will focus on the result for the focal variable - phase. While the full regression models will be reported as supplementary information, the coefficients for the other variables will not be interpreted (28).

We will estimate standardised effect sizes for all three outcomes (one primary and two secondary outcomes) by dividing the estimated coefficient $\boldsymbol{\beta_{01}}$ over the standard deviation of the dependent variable. Apart from providing estimates of practical significance of the findings, this has the additional benefit of allowing comparisons of effect sizes across the outcomes (which is not possible when comparing coefficients as the scales and ranges of the measures vary). While it is difficult to provide a general cut-off for what is a relevant effect size in population-based research, in this study anything above a 10% of a standard deviation change in continuous scores would be considered an effect with potentially practical significance at population level (29).

An additional approach to considering effect size as recommended for population-based research (29), is the population attributable fraction (PAF; 30). We will estimate the PAF only for the primary outcome if there is support for our hypothesis (i.e., significant impact of COVID-19 on depressive symptoms). This allows one to estimate the number of cases that are attributable to the exposure of interest (i.e., COVID-19) and hence from this we can estimate the proportion of cases fewer than might be expected in the absence of the COVID-19 pandemic. We will estimate this using the punaf package in Stata (31).

3.5.3. Sensitivity analyses

To describe the multivariate comparability of the two samples considering school- and individual-level variables, a propensity score will be estimated with a random intercept logistic regression model, first for individual level only and then for a model with both school- and individual-level baseline variables as predictors of "phase". We will separately visualise the distribution of the two propensity scores across the cohorts; and we will use the Stata module psmatch2 (-pstest-; 32) and 1-to-1 matching to report descriptive statistics for the included covariates (see Table 1), with and without matching. Although this will result in

some unmatched students across phases (due to unequal sample sizes), this descriptive method offers comprehensive insight into the comparability of the underlying samples with respect to the available characteristics.

While the applied mixed effects models can accommodate missing data in the dependent variable (under the MAR assumption), two sensitivity analyses will explore the sensitivity of our main finding by adding additional approaches to account for missing data. The first of these analyses will extend all models by inverse probability weighting for the probability to be a drop-out at follow-up (see Table 4 below and description of missing data analyses). The second of these analyses will use multiple imputation with chained equations using the full set of study variables as auxiliary variables to enhance the data for participants who are only partially observed at baseline. Results of the primary and secondary outcomes, including the two sensitivity analyses, will be reported as displayed in Table 5.

3.5.4 Exploratory analysis: subgroup differences

To examine whether the impacts of the pandemic were differently experienced by sub-groups of adolescents, we will conduct the main analysis with an interaction term between each modifier of interest and phase in a separate model for each modifier (gender, ethnicity, socio-economic status, and SEN). Given the modifiers of interest are all binary categorical (coded 0,1) these will be entered into the models as is, as these interaction terms remain directly interpretable. If an interaction term has a p-value <0.10 we will visualise the interaction results in a graph based on predicted margins from the model. A sensitivity analysis for these models will be conducted where all modifiers and their interaction with phase are included in the same model (33).

3.5.5. Missing data

We will first examine the rates and predictors of missing data at follow-up assessment, and these will be reported. Alongside the complete case analysis, we will 1) conduct weighted analysis to account for non-response at follow-up. Non-response weights will be created separately in each phase as the mechanisms generating missingness might vary between phases; and 2) use multiple imputation with chained equations (MICE) controlling for school clusters to complete data for participants who are partially observed at the baseline assessments (i.e., took part in the survey but did not complete the primary or secondary outcome measures).

3.5.6. Power analysis

The prospective evaluation of statistical power is constrained by the design of the original studies for which power analyses were published (21, 22). As in the main project, we assume for the primary outcome, self-reported depressive symptoms as measured by the SMFQ, a conservative school-level intraclass correlation of rho=0.10. Based on our current estimate of the database we expect around 185 analysed schools (phase 1, n = 96; phase 2, n = 89), and an average of 72.5 students per school (see Figure 2) and accepting a significance level of $p = .05$ and statistical power of $\beta = .80$, the minimally detectable effect size (MDES; 34) is estimated as MDES = .139 (all estimates obtained with Optimal Design; 34). Assuming potential additional losses on student-level of 10% due to inability to match data with NPD records increases this to MDES = .140.

For dichotomised SMFQ values, the analysis evaluates whether the share of students with a changed score differs between the two phases at follow-up. Based on estimates obtained with the same measure in the population-based Millennium Cohort Study (35), we assume a plausible range for the prevalence of increased levels of depressive symptoms before the pandemic is between .10 and .25. Expecting a pre-pandemic prevalence point estimate of .15, the prevalence after the pandemic would need to be either lower than .128 or above .174 to be detectable accepting a significance level of $p = .05$ and statistical power of $\beta = .80$ (.127 and .175 for 10% of student-level dropout). The addition of covariates with predictive power on any level (school or pupil), potentially further increases precision of estimates. The primary outcome analysis is well-powered to identify a potentially relevant effect of the COVID-19 pandemic and the societal response on pupils' depressive symptoms.

<Insert Table 2 here>

4. Results

<Insert Table 3 here>

<Insert Table 4 here>

<Insert Table 5 here>

5. Discussion

Limitations

The following will limit our ability to fully ascribe causation to our study findings (whether they support the hypothesis or not). First, the student composition of the two study phases might already differ at baseline. Second, differential attrition across the two phases: the response rates at follow-up for the two phases are unlikely to be the same, and the predictors of non-response might vary in the pre-pandemic and COVID-19 phases of the study. Third, there might be imbalances in the distribution of the interventions and controls across the two phases, and it is also plausible that differential effectiveness of the interventions across COVID-19 and pre-pandemic phases of the study might impact on our current analysis in unforeseeable ways. Finally, the pandemic experience could have had an impact on how students interpret or respond to the outcome measures.

6. Conclusion

Acknowledgments

We would also like to thank all the colleagues and researchers involved in the Education for Wellbeing Programme from which these data are drawn. We are grateful to all the schools and students who are part of this large study for their participation.

Ethical Statement

Ethical approval was obtained from University College London Research Ethics Committee [6735/009, 6735/014]. Informed consent was obtained from all participants.

Funding Statement

This work (the Education for Wellbeing Programme) was supported by the Department for Education, England [grant number: EOR/SBU/2017/015].

Data Accessibility

As per our trial protocols (21, 22), the raw data can only be shared after the trials are complete. An anonymised quantitative dataset was originally due to be made available in 2022 once the study had finished, however, due to disruption caused by the COVID-19 pandemic, this timeline has been delayed. Analysis code and the results log will be shared on the Open Science Framework when the full stage 2 paper is submitted/published.

Competing Interests

We have no competing interests. The views expressed are those of the authors and not necessarily those of the Department for Education, England or its arm's length bodies, or other Government Departments. The authors declare no conflict of interest and take sole responsibility for the content of this article.

Authors' Contributions

PP, JB, JD, and DH conceptualised the study with input from all authors. RM, JS, PP, and JB drafted the manuscript, and all authors inputted with critical revisions. JS is the data manager for the study, oversees the data management and planned NPD linkage and created the data flow diagram. DH was the Trials Manager for the study and oversaw data collection. RM, JS, and PP will conduct the analysis described in this registered report with supervisory input from JB.

References

1. Sadler K, Vizard T, Ford T, Marcheselli F, Pearce N, Mandalia D, et al. 2018. *Mental health of children and young people in England, 2017 summary of key findings*. Available at: https://files.digital.nhs.uk/A6/EA7D58/MHCYP%202017%20Summary.pdf
2. Fink E, Patalay P, Sharpe S, Holley S, Deighton J, Wolpert M. 2015. Mental health difficulties in early adolescence: A comparison of two cross-sectional studies in England from 2009 and 2014. *J Adolesc Health*, **56**(5), 502–507. (doi: 10.1016/j.jadohealth.2015.01.023)
3. Patalay P, Gage, SH. 2019. Changes in millennial adolescent mental health and health related behaviours over ten years: A population cohort comparison study. *Int J Epidemiol*, **48**(5), 1650–1664. (doi: 10.1093/ije/dyz006)
4. Viner R, Russell S, Saulle R, Croker H, Stansfeld C, Packer J, et al. 2021. Impacts of school closures on physical and mental health of children and young people: A systematic review. *medRxiv*. (doi:10.1101/2021.02.10.21251526)
5. Ford T, John A, Gunnell D. (2021). Mental health of children and young people during pandemic. *BMJ*, **372**(March), 1–2. (doi: 10.1136/bmj.n614)
6. Hafstad GS, Augusti E-M. 2021. A lost generation? COVID-19 and adolescent mental health. *Lancet Psychiatry*, **8**(8), 640–641. (doi: 10.1016/s2215-0366(21)00179-6)
7. Thombs BD, Bonardi O, Rice DB, Boruff JT, Azar M, He C, et al. 2020. A. Curating evidence on mental health during COVID-19: A living systematic review. *J Psychosom Res*, **133**, 110–113. (doi: 10.1016/j.jpsychores.2020.110113)
8. Achterberg M, Dobbelaar S, Boer OD, Crone EA. 2021. Perceived stress as mediator for longitudinal effects of the COVID-19 lockdown on wellbeing of parents and children. *Sci Rep*, **11**(1), 1–14. (doi: 10.1038/s41598-021-81720-8)
9. Collishaw S. 2015. Annual research review: Secular trends in child and adolescent mental health. *J Child Psychol Psychiatry Allied Discip*, **56**(3), 370–393. (doi: 10.1111/jcpp.12372)
10. Kim-Cohen J, Caspi A, Moffitt TE, Harrington H, Milne BJ, Poulton R. 2003. Prior juvenile diagnoses in adults with mental disorder. *Arch Gen Psychiatry*, **60**(7), 709. (doi:10.1001/archpsyc.60.7.709)
11. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. 2005. Lifetime Prevalence and Age-of-Onset distributions of of DSM-IV Disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*, **62**, 593–602. (doi: 10.1001/archpsyc.62.6.593)
12. Sun Y, Wu Y, Bonardi O, Krishnan A, He C, Boruff JT, et al. 2021. Comparison of Mental Health Symptoms prior to and during COVID-19: Evidence from a living Systematic Review and Meta-analysis. *medRxiv*. (doi: 10.1101/2021.05.10.21256920)
13. Magson NR, Freeman JYA, Rapee RM, Richardson CE, Oar EL, Fardouly J. 2021. Risk and protective factors for prospective changes in adolescent mental health during the COVID-19 Pandemic. *J Youth Adolesc*, **50**(1), 44–57. (doi: 10.1007/s10964-020-01332-9)
14. Ezpeleta L, Navarro JB, de la Osa N, Trepat E, Penelo E. 2020. Life conditions during COVID-19 lockdown and mental health in Spanish adolescents. *Int J Environ Res Public Health*, **17**(19), 1–13. (doi: 10.3390/ijerph17197327)
15. Xiang M, Yamamoto S, Mizoue T. 2020. Depressive symptoms in students during school closure due to COVID-19 in Shanghai. *Psychiatry Clin Neurosci*, **74**(12), 664–666. (doi: 10.1111/pcn.13161)
16. Thorisdottir IE, Asgeirsdottir BB, Kristjansson AL, Valdimarsdottir HB, Jonsdottir Tolgyes EM, Sigfusson J, et al. 2021. Depressive symptoms, mental wellbeing, and substance use among adolescents before and during the COVID-19 pandemic in Iceland: a longitudinal, population-based study. *Lancet Psychiatry*, **8**(8), 663–672. (doi: 10.1016/s2215-0366(21)00156-5)
17. Wright N, Hill J, Sharp H, Pickles A. 2021. Interplay between long-term vulnerability and new risk: Young adolescent and maternal mental health immediately before and during the COVID-19 pandemic. *JCPP Adv*, **1**(e12008). (doi: 10.1111/jcv2.12008)
18. Vizard T, Sadler K, Ford T, Newlove-Delgado T, McManus S, Marcheselli F, et al. 2020. *Mental Health of Children and Young People in England, 2020 Wave 1 follow up to the 2017 survey*, (July), 1–50. Available at: https://files.digital.nhs.uk/CB/C41981/mhcyp_2020_rep.pdf
19. NHS Digital. 2020. *Mental Health of Children and Young People in England, 2020 Wave 1 follow-up to the 2017 survey, survey design and methods report*. Available at: https://files.digital.nhs.uk/D1/D411D3/mhcyp_2020_meth.pdf
20. Creswell C, Shum A, Pearcey S, Skripkauskaite S, Patalay P, Waite P. Young people's mental health during the COVID-19 pandemic. *Lancet Child Adolesc Health*, **5**(8), 535–537. (doi: 10.1016/S2352-4642(21)00177-2)
21. Hayes D, Moore A, Stapley E, Humphrey N, Mansfield R, Santos J, et al. 2019. School-based intervention study examining approaches for well-being and mental health literacy of pupils in Year 9 in England: study protocol for a multischool, parallel group cluster randomised controlled trial (AWARE). *BMJ Open*, **9**(e029044). (doi: 10.1136/bmjopen-2019-029044)
22. Hayes D, Moore A, Stapley E, Humphrey N, Mansfield R, Santos J, et al. 2019. Promoting mental health and wellbeing in schools: Examining Mindfulness, Relaxation and Strategies for Safety and Wellbeing in English primary and secondary schools: Study protocol for a multi-school, cluster randomised controlled trial (INSPIRE). *Trials*, **20**(1), 4–8. (doi: 10.1186/s13063-019-3762-0)
23. Angold A, Costello EJ, Messer SC, Pickles A, Winder F, Silver D. 1995. The development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *Int J Methods Psychiatr Res*, **5**(4), 237–249.
24. Deighton J, Tymms P, Vostanis P, Belsky J, Fonagy P, Brown A, et al. 2013. The development of a school-based measure of child mental health. *J Psychoeduc Assess*, **31**(3), 247–257. (doi: 10.1177/0734282912465570)
25. Patalay P, Deighton J, Fonagy P, Vostanis P, Wolpert M. 2014. Clinical validity of the Me and My School questionnaire: A self-report mental health measure for children and adolescents. *Child Adolesc Psychiatry Ment Health*, **8**(1), 1–7. (doi: 10.1186/1753-2000-8-17)
26. Huebner ES. 1991. Initial development of the Student's Life Satisfaction Scale. *School Psychology International*, **12**(3), 231–240. (doi: 10.1177/0143034391123010)
27. Seligson JL, Huebner ES, Valois RF. 2003. Preliminary validation of the Brief Multidimensional Students' Life Satisfaction Scale (BMSLSS). *Soc Indic Res*, **61**(2), 121–145. (doi: 10.1023/A:1021326822957)
28. Westreich D, Greenland S. 2013. The table 2 fallacy: Presenting and interpreting confounder and modifier coefficients. *Am J Epidemiol*, **177**(4), 292–298. (doi: 10.1093/aje/kws412)
29. Matthay EC, Hagan E, Gottlieb LM, Tan ML, Vlahov D, Adler N, et al. 2021. Powering population health research: Considerations for plausible and actionable effect sizes. *SSM Popul Health*, **14**(100789). (doi: 10.1016/j.ssmph.2021.100789)
30. Mansournia MA, Altman, DG. 2018. Population attributable fraction. *BMJ (Clinical research ed.)*, **360**, k757. (10.1136/bmj.k757)
31. Newson R. 2010. PUNAF: Stata module to compute population attributable fractions for cohort studies. *Statistical Software Components* **S457193**, Boston College Department of Economics, revised 04 Sep 2015.
32. Leuven E, Sianesi B. 2003. PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. *Statistical Software Components* **S432001**, Boston College Department of Economics, revised 01 Feb 2018.
33. Keller MC. 2014. Gene x environment interaction studies have not properly controlled for potential confounders: The problem and the (simple) solution. *Biological Psychiatry*, **75**(1), 18–24. (doi: 10.1016/j.biopsych.2013.09.006)
34. Raudenbush SW, et al. 2011. *Optimal Design Software for multi-level and longitudinal research* (Version 3.01). Available from www.wtgrantfoundation.org
35. Patalay P, Fitzsimons E. 2017. *Mental ill-health among children of the new century: Trends across childhood with a focus on age 14*. Centre for Longitudinal Studies: London. Available from: https://www.ucl.ac.uk/ioe/sites/ioe/files/patal

ay_fitzsimons_mental_ill-
health_among_children_of_the_new_century_
-_september_2017.pdf

Tables

Table 1. Covariate variables used in the analysis, data source, scoring, and role in the analysis

Variable	Source	Scoring	Role
School level			
Phase	Based on phase of recruitment	0 = Phase 1 (pre-COVID-19 group) 1 = Phase 2 (COVID-19 group)	Focal variable
Urban / rural	GIAS service, previously known as Edubase	0 = Town/rural 1 = City	Covariate
School deprivation - free school meal eligibility	GIAS service, previously known as Edubase	% of pupils in school	Covariate
Prior interventions (baseline school mental health provision)	Baseline school mental health provision survey	0 = No mental health intervention/support 1 = Prior interventions structured lessons/other mental health support	Covariate
Individual level			
Free school meal (FSM) eligibility	NPD – codes EVERFSM_ALL/EVERFSM_6	0 = Not eligible 1 = Eligible	Covariate (and modifier)
Gender	Direct from schools	0 = Male 1 = Female	Covariate (and modifier)
Special Educational Needs (SEN)	NPD - code SENprovisionMajor	0 = No SEN 1 = SEN	Covariate (and modifier)
Ethnicity	NPD – code EthnicGroupMajor	0 = White 1 = Non-white ethnic minority	Covariate (and modifier)
Age (year group)	Direct from schools	Year 7, 8 or 9	Covariate
Baseline mental health outcome score	From baseline survey	Centred score	Covariate

Note. *NPD = National Pupil Database, GIAS = Get Information About Schools service*

Table 2. Study design summary

Research Questions	Hypotheses	Outcome Measures	Sampling Plan (N, power analyses)	Analysis Plan	Rationale for deciding the sensitivity of the test for confirming or disconfirming the hypothesis	Interpretation given different outcomes	Theory that could be disproved by the outcomes
What impact has the COVID-19 pandemic had on adolescent mental health, specifically, depressive symptoms (primary outcome), externalising difficulties, and life satisfaction (secondary outcomes)?	PRIMARY OUTCOME, H1: after controlling for baseline variables, levels in depressive symptoms will be higher during the COVID-19 pandemic compared to before; SECONDARY OUTCOME, H2: after controlling for baseline variables, levels in externalising difficulties will be higher during the COVID-19 pandemic compared to before;	Short Mood and Feelings Questionnaire (23) The continuous depressive symptoms score will be the primary outcome measure. A binary score to assess impact on prevalence of case-ness will be generated using the established cut off score (≥ 12) indicating high levels of depressive symptoms (23). Behavioural difficulties subscale of the Me and My Feelings questionnaire (24, 25)	The prospective evaluation of statistical power is constrained by the design of the original studies for which power analyses were published (21, 22). As in the main project, we assume for the primary outcome, self-reported depressive symptoms as measured by the SMFQ, a conservative school-level intraclass correlation of rho=0.10. Based on our current estimate of the database we expect around 185 analysed schools (phase 1, N=96; phase 2, n=89), and an average of 72.5 students per school (see Figure 2.), and accepting a significance level of $p = .05$ and statistical power of $\beta = .80$, the minimally detectable effect size is estimated as MDES = .139 (all estimates obtained with Optimal Design; (29)). Assuming potential additional losses on student-level of 10% due to inability to match data with NPD records increases this to MDES =	All analyses will be conducted using Stata software. The primary outcome analysis will use a random intercept (for schools) linear multivariable regression model with phase (exposure to COVID-19 pandemic) included as key predictor and all individual and school-level variables as covariates. For the dichotomous high depressive symptoms and externalising symptoms, we will run similar random	While the applied mixed effects models can accommodate missing data in the dependent variable (under the MAR assumption), two sensitivity analyses will explore the sensitivity of our main finding by adding additional approaches to account for missing data. The first of these analyses will extend all models by inverse probability	The primary result will be the coefficient for "Phase", which if found statistically significant at $p < .05$, will be interpreted as the COVID-19 pandemic having potentially had an effect on adolescent depressive symptoms. The direction of the potential effect is "increased adolescent depressive symptoms" if the SMFQ score is higher in phase 2	There is no specific theory that we will be able to disprove. However, if no significant effect is found for the impact of the COVID-19 pandemic on adolescent mental health outcomes, or if the significant effect is in the opposite direction indicating improvements to adolescent mental health, then the negative impact of the pandemic on adolescent

	SECONDARY OUTCOME, H3: after controlling for baseline variables, levels in life satisfaction will be lower during the COVID-19 pandemic compared to before;	Responses will be summed to create a total behavioural difficulties score, with higher scores indicating greater difficulties. The scale has an established cut-off score of ≥6, which will be used for analysis examining a binary outcome of high externalising difficulties (24). Huebner Life Satisfaction Scale (LSS) (26, 27) A total score is created summing responses from the 7 items, with higher total scores indicating greater life satisfaction.	.140. For dichotomised SMFQ values, the analysis evaluates whether the share of students with a changed score differs between the two phases at follow-up. Based on estimates obtained with the same measure in the population-based Millennium Cohort Study (30), we assume a plausible range for the prevalence of increased levels of depressive symptoms before the pandemic is between .10 and .25. Expecting a pre-pandemic prevalence point estimate of .15, the prevalence after the pandemic would need to be either lower than .128 or above .174 to be detectable accepting a significance level of $p = .05$ and statistical power of $\beta = .80$ (.127 and .175 for 10% of student-level dropout). Both would usually be considered small effect sizes in the field of research and practice. The addition of covariates with predictive power on any level (school or pupil), potentially further increases precision of estimates. The primary outcome analysis is well-powered to identify a potentially relevant effect of the COVID-19 pandemic and the societal response on pupils' depressive symptoms.	intercept logistic multivariable regression models and report odds ratios. The same strategy will be employed for the secondary outcomes, but we will be clear that the primary outcome analysis takes priority in interpretation of the results. As the selection of independent variables is determined by the variables available in the main study, the reporting will focus on the result for the focal variable - phase. While the full regression models will be reported as supplementary information, the coefficients for the other variables will not be interpreted (28).	weighting for the probability to be a drop-out at follow-up. The second of these analyses will use multiple imputation with chained equations using the full set of study variables as auxiliary variables to enhance the data for participants who are only partially observed at baseline.	(and "decreased" if higher in phase 1). If the coefficient is not statistically significant, the primary outcome analysis will be interpreted as "no supporting evidence for a difference was found".	mental health will be called into question.

Were there socio-demographic differences (based on gender, ethnicity, socio-economic disadvantage, and special educational needs) in the impact of the COVID-19 pandemic on adolescent mental health outcomes?	EXPLORATORY ANALYSIS, H4: Due to the exploratory nature no hypotheses are made.	Gender (male or female), socio-economic position assessed using eligibility for free school meals (FSM eligible or not), ethnicity (white or ethnic minority) and special education needs status (SEN, yes or no). Information on pupil-level socio-demographic characteristics will be available via linkage to the National Pupil Database (NPD).		To examine whether the impacts of the pandemic were differently experienced by sub-groups of adolescents, we will conduct the main analysis with an interaction term between each modifier of interest and phase in a separate model for each modifier (gender, ethnicity, socio-economic status, and SEN).	A sensitivity analysis for these models will be conducted where all modifiers and their interaction with phase are included in the same model.	If an interaction term has a p-value <0.10 we will visualise the interaction results in a graph based on predicted margins from the model.	If no socio-demographic differences are found for the impact of the COVID-19 pandemic on adolescent mental health, the current study will call into question findings from existing studies that do not include pre-pandemic assessments of mental health.

Table 3. Example of potential table presenting descriptive statistics (count, %) of school-level characteristics for each phase at baseline and 1-year follow-up

School Characteristics	School Count (%)			
	Phase 1		Phase 2	
	Pre-COVID-19 group		COVID-19 group	
	Baseline	Follow-up	Baseline	Follow-up
School allocation				
Control	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
Intervention	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
FSM eligibility				
Bottom third	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
Middle third	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
Upper third	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
Geographic location				
London	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
Manchester	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
Bath & Bristol	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
Durham	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
School location				
City	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
Town/Rural	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
Prior Interventions				
Structured lessons/Other mental health support	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
No mental health intervention/support	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)

Note. *In phase 1, the category 'Lost during one-year follow-up' refers to schools that left the EfW programme. In phase 2, this category also refers to schools that, despite remaining in the EfW programme, were unable to complete pupil surveys.*

Table 4. Example of potential result table presenting coefficients (coeff; for continuous outcomes) and odds ratios (OR; for binary outcomes) and their 95% Confidence Intervals (95% CI)

	Depressive symptoms *(primary outcome)* coeff (95% CI)	High depressive symptoms (binary) OR (95% CI)	Externalising coeff (95% CI)	High externalising (binary) OR (95% CI)	Life satisfaction coeff (95% CI)
Covariate-adjusted	x.xx (xx-xx)	x.xx (xx-xx)	x.xx (xx-xx)	x.xx (xx-xx)	x.xx (xx-xx)
Covariate-adjusted + IPW for drop-out	x.xx (xx-xx)	x.xx (xx-xx)	x.xx (xx-xx)	x.xx (xx-xx)	x.xx (xx-xx)
Covariate-adjusted + IPW for drop-out + MI for observed missing	x.xx (xx-xx)	x.xx (xx-xx)	x.xx (xx-xx)	x.xx (xx-xx)	x.xx (xx-xx)

Note. *All models adjusted for school-level and individual-level variables listed in Table 1; IPW = inverse probability weight; MI = multiple imputation, fully conditional.*

Table 5. Example of potential result table presenting results from the modification analyses by individual gender, ethnicity, FSM and SEN for coefficients (coeff; for continuous outcomes) and odds ratios (OR; for binary outcomes) and their 95% Confidence Intervals (95% CI).

	Depressive symptoms *(primary outcome)* coeff, (95% CI), p-value	High depressive symptoms (binary) OR, (95% CI), p-value	Externalising coeff, (95% CI), p-value	High externalising (binary) OR, (95% CI), p-value	Life satisfaction coeff, (95% CI), p-value
Phase x Gender (female)	x.xx	x.xx	x.xx	x.xx	x.xx
Phase x SEP (FSM eligible)	x.xx	x.xx	x.xx	x.xx	x.xx
Phase x Ethnicity (ethnic minority)	x.xx	x.xx	x.xx	x.xx	x.xx
Phase x SEN (SEN - yes)	x.xx	x.xx	x.xx	x.xx	x.xx

Note. *The coefficient reported is for the interaction term between phase and each modifier and the p-value. The results from any model indicating the presence of effect modification will be described in a visualisation. All models adjusted for school-level and individual-level variables listed in Table 1; IPW = inverse probability weight; MI = multiple imputation, fully conditional*

Figures

Figure 1. Study design: process chart with timelines and assessments in each phase

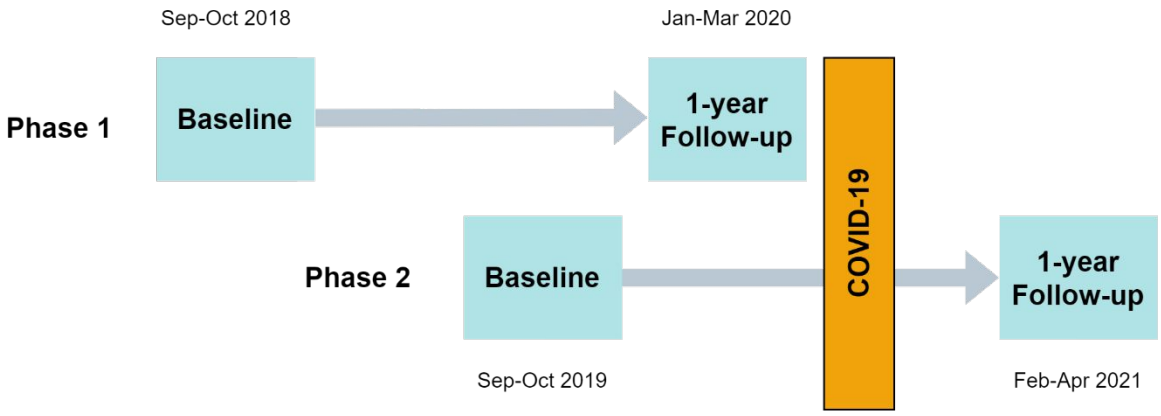

Figure 2. Flowchart of the current study - natural COVID-19 experiment

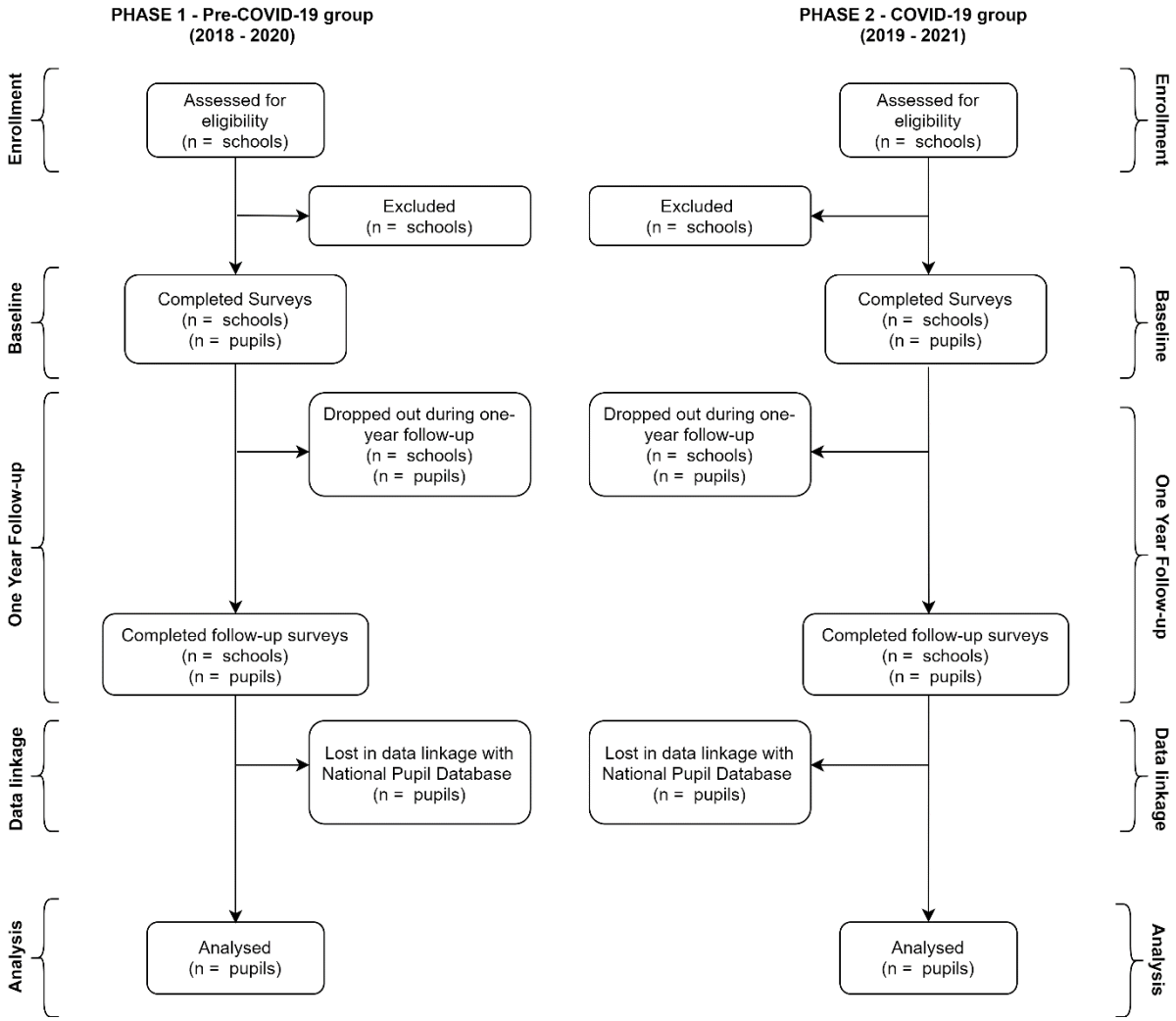

Figure and table captions

Figure 1. Study design: process chart with timelines and assessments in each phase

Figure 2. Flowchart of the current study - natural COVID-19 experiment

Table 1. Covariate variables used in the analysis, data source, scoring, and role in the analysis

Table 2. Study design summary

Table 3. Example of potential table presenting descriptive statistics (count, %) of school-level characteristics for each phase at baseline and 1-year follow-up

Table 4. Example of potential result table presenting coefficients (coeff; for continuous outcomes) and odds ratios (OR; for binary outcomes) and their 95% Confidence Intervals (95% CI)

Table 5. Example of potential result table presenting results from the modification analyses by individual gender, ethnicity, FSM and SEN for coefficients (coeff; for continuous outcomes) and odds ratios (OR; for binary outcomes) and their 95% Confidence Intervals (95% CI).