BMJ Open

# Pre-COVID-19 pandemic health-related behaviours in children (2018–2020) and association with being tested for SARS-CoV-2 and testing positive for SARS-CoV-2 (2020–2021): a retrospective cohort study using survey data linked with routine health data in Wales, UK

Emily Marchant ,[1,2] Emily Lowthian ,[1,2] Tom Crick ,[2] Lucy J Griffiths ,[1] Richard Fry ,[1,3] Kevin Dadaczynski,[4,5] Orkan Okan,[6] Michaela James ,[1,3] Laura Cowley,[1] Fatemeh Torabi,[1] Jonathan Kennedy ,[1,3] Ashley Akbari ,[1] Ronan Lyons ,[1] Sinead Brophy [1,3]

**To cite:** Marchant E, Lowthian E, Crick T, *et al*. Pre-COVID-19 pandemic health-related behaviours in children (2018–2020) and association with being tested for SARS-CoV-2 and testing positive for SARS-CoV-2 (2020–2021): a retrospective cohort study using survey data linked with routine health data in Wales, UK. *BMJ Open* 2022;**12**:e061344. doi:10.1136/bmjopen-2022-061344

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2022-061344).

Received 25 January 2022
Accepted 07 August 2022



For numbered affiliations see end of article.

**Correspondence to**
Dr Emily Marchant;
e.k.marchant@swansea.ac.uk

## ABSTRACT

**Objectives** Examine if pre-COVID-19 pandemic (prior March 2020) health-related behaviours during primary school are associated with (1) being tested for SARS-CoV-2 and (2) testing positive between 1 March 2020 and 31 August 2021.

**Design** Retrospective cohort study using an online cohort survey (January 2018 to February 2020) linked with routine PCR SARS-CoV-2 test results.

**Setting** Children attending primary schools in Wales (2018–2020), UK, who were part of the Health and Attainment of Pupils in a Primary Education Network (HAPPEN)_school network.

**Participants** Complete linked records of eligible participants were obtained for n=7062 individuals. 39.1% (n=2764) were tested (age 10.6±0.9; 48.9% girls) and 8.1% (n=569) tested positive for SARS-CoV-2 (age 10.6±1.0; 54.5% girls).

**Main outcome measures** Logistic regression of health-related behaviours and demographics were used to determine the ORs of factors associated with (1) being tested for SARS-CoV-2 and (2) testing positive for SARS-CoV-2.

**Results** Consuming sugary snacks (1–2 days/week OR=1.24, 95% CI 1.04 to 1.49; 5–6 days/week OR=1.31, 95% CI 1.07 to 1.61; reference 0 days), can swim 25 m (OR=1.21, 95% CI 1.06 to 1.39) and age (OR=1.25, 95% CI 1.16 to 1.35) were associated with an increased likelihood of being tested for SARS-CoV-2. Eating breakfast (OR=1.52, 95% CI 1.01 to 2.27), weekly physical activity ≥60 min (1–2 days OR=1.69, 95% CI 1.04 to 2.74; 3–4 days OR=1.76, 95% CI 1.10 to 2.82; reference 0 days), out-of-school club participation (OR=1.06, 95% CI 1.02 to 1.10), can ride a bike (OR=1.39, 95% CI 1.00 to 1.93), age (OR=1.16, 95% CI 1.05 to 1.28) and girls (OR=1.21, 95% CI 1.00 to 1.46) were associated with an increased likelihood of testing positive for SARS-CoV-2. Living in least deprived areas (quintile 4 OR=0.64, 95% CI 0.46 to 0.90; quintile 5 OR=0.64, 95% CI 0.46 to 0.89) compared with the most deprived (quintile 1) was associated with a decreased likelihood.

**Conclusions** Associations may be related to parental health literacy and monitoring behaviours. Physically active behaviours may include coparticipation with others and exposure to SARS-CoV-2. A risk-versus-benefit approach must be considered in relation to promoting these health

## STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ Investigation of the association of prepandemic child health-related behaviour measures with subsequent SARS-CoV-2 testing and infection.
⇒ Reporting of multiple child health behaviours linked at an individual level to routine records of SARS-CoV-2 testing data through the Secure Anonymised Information Linkage Databank, using complete case analysis.
⇒ Child-reported health behaviours were measured before the COVID-19 pandemic (1 January 2018 to 28 February 2020) which may not reflect behaviours during COVID-19.
⇒ Health behaviours captured through the national-scale Health and Attainment of Pupils in a Primary Education Network (*HAPPEN*) survey represent children attending schools that engaged with the *HAPPEN* Wales primary school network and may not be representative of the whole population of Wales.
⇒ The period of study for PCR testing includes a time frame with varying prevalence rates, approaches to testing children (targeted and mass testing) and restrictions which were not measured in this study.

behaviours, given the importance of health-related behaviours such as childhood physical activity for development.

## BACKGROUND

The COVID-19 pandemic caused by SARS-CoV-2 has resulted in widespread disruption to the lives of children across the world, and has contributed to widened inequalities in children's health, well-being and education.[1 2] Childhood is a critical developmental period during which health behaviours are established which transcend into adolescence and adulthood.[3] The Organisation for Economic Co-operation and Development (OECD) recognised current trends in children's health, highlighting typical health behaviours of school-age children that warrant further research in order to better design policies that improve children's health outcomes.[4 5] These include nutrition-related behaviours such as fruit and vegetable intake, consumption of sugary foods and breakfast consumption, physical activity and sedentary behaviours and sleep. The establishment of these health behaviours during childhood is highly influenced by parental mechanisms and monitoring behaviours, particularly in children aged under 12.[6–8]

While evidence has demonstrated the negative impact of the COVID-19 pandemic on children's health-related behaviours including reduced physical activity, increased sedentary behaviour and poorer nutrition,[1 9] it is unclear if this association is bidirectional. That is, whether these health behaviours are associated with likelihood of SARS-CoV-2 infection. Within the adult population, emerging evidence suggests a plausible relationship between prepandemic health risk behaviours such as physical inactivity and poor nutrition with SARS-CoV-2 infection and severity of disease,[10–13] and increased risk of other infectious diseases.[14] This is attributed to the important role health behaviours play in shaping cardiometabolic health and immune system function. Indeed, research shows links to the early years including critical early developmental stages with subsequent risk of developing chronic inflammation, which is associated with non-communicable disease risk and mortality during adulthood.[15] Health behaviours such as adequate nutrient intake[16] and physical activity[17] are required for the regulation and function of the immune system.

As a result, researchers have advocated for consideration to be placed on the role of these health behaviours in future endemic/pandemic scenarios.[17] However, research to date has concentrated on adults, explored single health behaviours or examined those with severe COVID-19 infection and hospitalisation.[18 19] The focus of research within the childhood population has principally been placed on clinical outcomes as opposed to lifestyle outcomes, including identifying the clinical characteristics of severe infection, the presence of comorbidities, common symptoms such as cough and clinical biomarkers.[20 21] While serious COVID-19 illness in children is relatively rare, mild or asymptomatic infection is common.[22] Positive SARS-CoV-2 tests require periods of self-isolation, impacting children's physical health and well-being, limiting opportunities for children to engage in health-promoting behaviours essential for optimal development such as regular physical activity.[9 23] Therefore, research examining the role of these health behaviours in a childhood population within the context of the COVID-19 pandemic is warranted.

Identifying the prepandemic health-related behavioural characteristics of children requiring a SARS-CoV-2 test or testing positive for SARS-CoV-2 infection and hypothesising potential mechanisms through which these may operate, including exposures, sociodemographic and parental influences, could yield insight to inform the current COVID-19 pandemic and future pandemic/endemic scenarios. This can also allow targeted intervention to minimise transmission risk that complements national public health measures and guidelines, and importantly, mitigates the disruption to children's lives, and prevent further exacerbation of pre-existing inequalities, safeguarding their health, well-being and education.

In Wales (one of the four nations of the UK, with devolved health and social care policies), approaches to performing PCR tests on children during the period of study included the presence of COVID-19 symptoms, if identified as a close contact to a positive case (eg, household contacts), or as a follow-up PCR test as encouraged in guidance at the time following a positive lateral flow test (LFT) (eg, showing symptoms or a close contact and having a positive LFT performed in the home).[24] Uptake of testing within the childhood population requires parental monitoring behaviours; for example, providing transport to testing facilities and parental health literacy through identification of symptoms.

This study investigates the association of prepandemic (prior to 1 March 2020) health-related behaviours self-reported by children aged 8–11 years during primary school before the COVID-19 pandemic between 1 January 2018 and 28 February 2020, with two outcomes: the odds of ever having a SARS-CoV-2 PCR test and the odds of ever testing positive for SARS-CoV-2 during the period of study. We aim to examine whether these self-reported markers of health-related behaviours reported before pandemic are associated with the likelihood of: (1) ever being tested for SARS-CoV-2 and (2) ever testing positive for SARS-CoV-2 between 1 March 2020 and 31 August 2021.

## METHODS

### Study design

This retrospective cohort study was conducted through the Health and Attainment of Pupils in a Primary Education Network (*HAPPEN*) primary school network.[25] *HAPPEN* was established in Wales, UK in 2014, following research with head teachers who advocated for increased collaboration to prioritise pupils' health and well-being,[26 27] and is a platform for conducting school-based

research.[2 28–30] The network brings together primary schools with research and runs up to the current date. School participation in *HAPPEN* is voluntary and is either once, annually or biannually (eg, to evaluate school-based interventions). Through *HAPPEN*, children aged 8–11 (years 4–6) complete the *HAPPEN* survey, an online cohort survey that captures a range of validated self-reported health behaviours including physical activity, nutrition and sleep.[31] Retrospective health-related behaviour data were obtained from responses from the *HAPPEN* survey completed before pandemic between 1 January 2018 and 28 February 2020.

These retrospective survey responses were linked with PCR SARS-CoV-2 test results obtained from the Pathology COVID-19 Daily (PATD) routine data set between 1 March 2020 and 31 August 2021. The PATD data set contains pillar 1 (swab testing in Public Health England labs, National Health Service (NHS) Wales labs and NHS hospitals for those with a clinical need, and health and care workers) and pillar 2 (swab testing for the wider population, as set out in government guidance) individual results from PCR tests (negative (suspected), positive (confirmed) for SARS-CoV-2).[32] The period of interest (1 March 2020 to 31 August 2021) includes a time frame of varying approaches to testing children, documented in timeline format in online supplemental appendix 1.[32] This includes targeted (ie, symptomatic and suspected positive cases, identified as a close contact of a positive case) and mass testing (ie, between February 2021 and April 2021, the use of LFTs in the school setting for pupils aged 11 and above (secondary school age) to identify asymptomatic positive cases, with guidance for positive LFTs encouraging follow-up PCR tests).

Linkage was performed using the *Secure Anonymised Information Linkage* (SAIL) Databank.[33–35] Data were linked at the individual level using an anonymous linkage field (ALF) to identify participants and link SARS-CoV-2 test results (figure 1). The REporting of studies Conducted using Observational Routinely-collected Data (RECORD) checklist[36] for this study is presented in online supplemental appendix 2.

## Ethics

Electronic data (survey responses) were stored in secure files only accessible to the research team. The routine data used in this study are available in the SAIL Databank and are subject to review by an independent Information Governance Review Panel (IGRP), to ensure proper and appropriate use of SAIL data. Before any data can be accessed, approval must be received from the IGRP. When access has been approved, it is accessed through a privacy-protecting safe haven and remote access system referred to as the SAIL Gateway. SAIL has established an application process to be followed by anyone who would like to access data. This study has been approved by the SAIL IGRP (project reference: 0911).

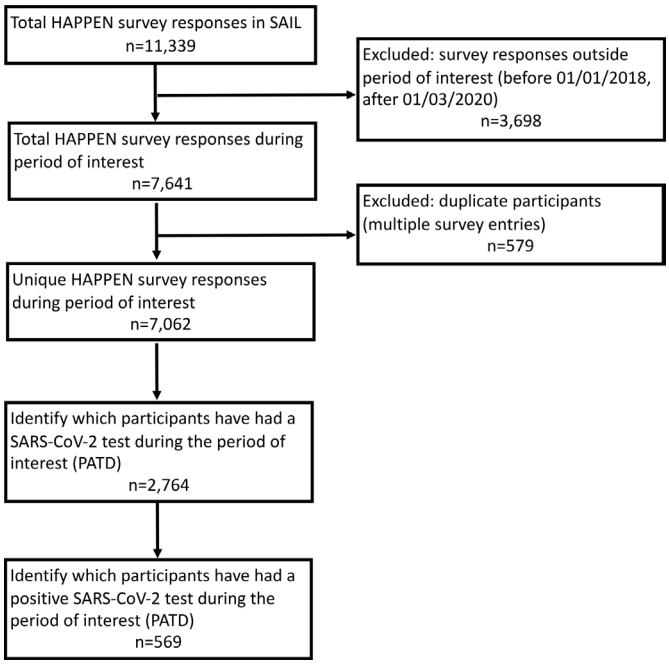



**Figure 1** Cohort flow diagram. HAPPEN, Health and Attainment of Pupils in a Primary Education Network; PATD, Pathology COVID-19 Daily; SAIL, Secure Anonymised Information Linkage.

## The *HAPPEN* survey and linked SAIL data

All primary schools (n=1203) in Wales, UK were invited to participate in the *HAPPEN* survey between 1 April 2014 and 28 February 2020 via a number of methods including email, social media promotion and through stakeholders in health and education (including local authority health and well-being teams, regional education consortia). Prior to 2018, *HAPPEN* was established in three of the local authorities (total n=22) in Wales. From 2018 to the period of interest, *HAPPEN* began its expansion to primary schools across Wales. Between 1 January 2018 and 28 February 2020, there were n=305 primary schools registered with *HAPPEN* (25% of primary schools in Wales). Participating in *HAPPEN* is voluntary and this study comprises a convenience sample of children attending n=129 primary schools (representing a 42% response rate of registered *HAPPEN* primary schools) from 16 out of 22 local authorities that participated in the *HAPPEN* survey during the period of interest (1 January 2018 and 28 February 2020). Schools were invited to share details of the survey with parents/guardians (including information sheets). To participate in the *HAPPEN* survey and link data to routine records, child assent was required in addition to parental consent (between 2014 and 2018) and opt-out parental consent (2019 onwards).

The *HAPPEN* survey is completed by children aged 8–11 as a self-guided activity within the school setting as a classroom activity with supervision from a teacher/teaching assistant. The survey takes approximately 30 min to complete and includes validated self-report measures of typical health behaviours including physical activity, screen time, nutrition, sleep and well-being.[31] A

full copy of the survey can be found in online supplemental appendix 3, and items, response categories and the coding framework included within analyses in online supplemental appendix 4.

The process of data coding involved two researchers. The first (MJ) cleaned the raw data (including checking for duplicate entries), removed identifiable information and generated a unique participant ID number to protect participants' anonymity. The second (EM) researcher coded the anonymised raw data set using STATA (V.16) to produce a data set for analyses. This *HAPPEN* data set was uploaded to the SAIL Databank, a trusted research environment containing individual-level anonymised population-scale data sources about the population of Wales, that enables secure data linkage and analysis for research, to be linked with SARS-CoV-2 testing data from the PATD data set. To link the data, the person-based identifiable data are separated from the survey data and sent to a trusted third party, Digital Health and Care Wales (the national organisation that designs and builds digital services for health and social care in Wales). The survey data are sent to SAIL using a secure file upload. A unique anonymous linkage field (ALF) is assigned to the person-based record before it is joined to clinical data via a system linking field. The ALF was used to link records at the individual level between the *HAPPEN* data set and PATD data set containing PCR testing data. This data set was accessible to authors listed from the Population Data Science group, Swansea University.

### Quantitative analysis

The primary outcomes were (1) whether the child was ever PCR tested for the SARS-CoV-2 virus and (2) whether the child had any positive SARS-CoV-2 test between 1 March 2020 and 31 August 2021. Participants were assigned a binary code for any SARS-CoV-2 test during the period of interest (1: PCR tested at least once for SARS-CoV-2 between 1 March 2020 and 31 August 2021; 0: no PCR SARS-CoV-2 test) and again for any positive SARS-CoV-2 test during the period of interest (1: any positive SARS-CoV-2 PCR test between 1 March 2020 and 31 August 2021; 0: negative PCR test for SARS-CoV-2; 0: not PCR tested for SARS-CoV-2 (unknown)). In the case of multiple PCR tests, the first occurrence was used. Participants were assumed to have remained in Wales during the period of interest. Eligibility criteria (see cohort flow diagram, figure 1) within final analyses models were any unique participant with complete linked survey and routine records. Inclusion dates of survey responses for analyses were between 1 January 2018 and 28 February 2020. Complete case multivariable logistic regression analyses, adjusting for confounding variables (sex, age on 1 March 2020, area-level deprivation using the Welsh Index of Multiple Deprivation (WIMD)[37] (version 2019)) and clustered by school (using sandwich estimator to account for differences between schools), determined the Odds Ratios (OR) for (1) ever being PCR tested for SARS-CoV-2 virus and (2) ever having a positive PCR SARS-CoV-2 test during the period of interest. Missing categories of data (sex and WIMD data obtained through the SAIL Databank) were tested to see if they significantly predicted any outcomes.

Independent variables as measures of typical prepandemic health-related behaviours included within analyses were obtained retrospectively from the *HAPPEN* survey, completed between 1 January 2018 and 28 February 2020 (online supplemental appendix 4). Health-related behaviour measures included in multivariable analyses are recognised by the OECD as typical health behaviour trends during childhood that warrant research.[4 5] These related to the behaviours from the previous day (ate breakfast, travel actively to and/or from school, number of fruit/vegetable portions consumed, number of times teeth brushed, hours of sleep), frequency of behaviours every day, the previous 7 days (physically active ≥60 min, sedentary/screen time ≥2 hours, felt tired, ate a sugary snack) and general items including participation in number of out-of-school clubs, can ride a bike and can swim 25 m. A list of variables included in analyses, coding response categories and coding framework is presented in online supplemental appendix 4. Independent variables were entered concurrently and examined for association with the outcomes (1) ever PCR tested for SARS-CoV-2 and (2) ever tested positive for SARS-CoV-2 between 1 March 2020 and 31 August 2021.

### Patient and public involvement

The SAIL Databank has a Consumer Panel that provides the public's perspective on data linkage research. The Panel members are involved in all elements of the SAIL Databank process, from developing ideas, advising on bids through approval processes (via the independent IGRP), to disseminating research findings. For more information visit https://saildatabank.com/about-us/public-engagement/.

## RESULTS

Survey responses were obtained from n=11 339 participants (figure 1). Survey responses outside the period of interest (before 1 January 2018 and after 28 February 2020) were excluded (n=3698), followed by duplicate participants (occasions of multiple survey entries, n=579). In the case of duplicates, the most recent instance of survey participation was used. Complete linked unique records of participants meeting eligibility criteria were obtained for n=7062 individuals. Table 1 presents the descriptive statistics of the study sample by ever PCR tested for SARS-CoV-2 and ever tested positive for SARS-CoV-2 between 1 March 2020 and 31 August 2021. Of the total sample, 39.1% (n=2764) were PCR tested for SARS-CoV-2 and 8.1% (n=569) tested positive for SARS-CoV-2. The mean age on 1 March 2020 (start of period of interest) was 10.6 (±0.9) for those PCR tested (table 1) and 10.6 (±1.0) for those tested positive for SARS-CoV-2 (table 2). The time between the *HAPPEN* survey date and the SARS-CoV-2

**Table 1** Descriptive statistics of study sample by PCR tested for SARS-CoV-2 and PCR test positive for SARS-CoV-2 between 1 March 2020 and 31 August 2021

| | Tested for SARS-CoV-2 % (n) | Not tested for SARS-CoV-2 % (n) | Tested positive for SARS-CoV-2 % (n) | Tested negative/not tested (unknown) for SARS-CoV-2 % (n) |
|---|---|---|---|---|
| Sample | 39.1 (2764) | 60.9 (4298) | 8.1 (569) | 91.9 (6493) |
| Age at time of *HAPPEN* survey | 10.1±0.8 | 9.9±0.9 | 10.1±0.8 | 9.9±0.8 |
| Age on 1 March 2020 (start of period of interest) | 10.6±0.9 | 10.3±1.1 | 10.6±1.0 | 10.4±1.0 |
| Number of days between *HAPPEN* survey and SARS-CoV-2 test (median (IQR)) | 588 (385–685) | | 672 (599–715) | |
| Sex | | | | |
| Boy | 49.3 (1363) | 46.7 (2005) | 44.3 (252) | 48.0 (3116) |
| Girl | 48.9 (1352) | 51.8 (2226) | 54.5 (310) | 50.3 (3268) |
| Missing | 1.8 (49) | 1.5 (67) | 1.2 (7) | 1.7 (109) |
| WIMD 2019 quintiles | | | | |
| 1 (most deprived) | 24.3 (672) | 23.9 (1,025) | 28.5 (162) | 23.6 (1535) |
| 2 | 19.9 (551) | 19.02 (826) | 19.7 (112) | 19.5 (1265) |
| 3 | 16.5 (455) | 17.4 (748) | 17.6 (100) | 17.0 (1103) |
| 4 | 15.6 (431) | 15.8 (678) | 14.1 (80) | 15.9 (1029) |
| 5 (least deprived) | 18.0 (497) | 16.8 (771) | 16.5 (94) | 17.3 (1124) |
| Missing | 5.7 (158) | 7.0 (300) | 3.7 (21) | 6.7 (437) |

See online supplemental appendix 4 for variable codebook. Full descriptive statistics table is presented in online supplemental appendix 5.
HAPPEN, Health and Attainment of Pupils in a Primary Education Network; WIMD, Welsh Index of Multiple Deprivation.

PCR test date (median number of days (IQR)) was 588 (385–685) days for being PCR tested and 672 (599–715) days for PCR testing positive for SARS-CoV-2. Complete case analyses are presented. The maximum missing data were 7% (see table 1). We tested if missing categories of data (sex and WIMD obtained through the SAIL Databank) significantly predicted any outcomes and found that no missing categories significantly predicted the outcomes. Therefore, missing data were assumed to be at random through data linkage.[38] Unadjusted multivariable logistic regression analyses are presented in online supplemental appendix 5.

Table 2 presents the multivariable logistic regression for children ever PCR tested for SARS-CoV-2 between 1 March 2020 and 31 August 2021. The model showed a low goodness of fit ($R^2$=0.02). Children who reported to eat breakfast (OR=1.16, 95% CI 0.99 to 1.36, reference: did not eat breakfast, p<0.1), consume sugary snacks on 1–2 days (OR=1.24, 95% CI 1.04 to 1.49) and 5–6 days (OR=1.31, 95% CI 1.07 to 1.61) compared with 0 days, participate in more out-of-school clubs (OR=1.02, 95% CI 1.00 to 1.04), able to ride a bike (OR=1.15, 95% CI 0.98 to 11.35, reference: cannot ride a bike, p<0.1) and able to swim 25 m (OR=1.21, 95% CI 1.06 to 1.39, reference: cannot swim 25 m) were more likely to be PCR tested for SARS-CoV-2. Older children (OR=1.25, 95% CI 1.16 to 1.35) were also more likely to be PCR tested for SARS-CoV-2, and compared with quintile 1 (most deprived), those in WIMD quintiles 3 (OR=0.85, 95% CI 0.70 to 1.03, p<0.1) and 5 (OR=0.85, 95% CI 0.72 to 1.02, p<0.1) were less likely to be PCR tested for SARS-CoV-2. Unadjusted multivariable logistic regression analyses are presented in online supplemental appendix 6.

Table 3 presents the multivariable logistic regression for children ever PCR tested positive for SARS-CoV-2 between 1 March 2020 and 31 August 2021. Children were more likely to test positive for SARS-CoV-2 if they reported to eat breakfast (OR=1.52, 95% CI 1.01 to 2.27, reference: did not eat breakfast), be physically active for ≥60 min on 1–2 days (OR=1.69, 95% CI 1.04 to 2.74), 3–4 days (OR=1.76, 95% CI 1.10 to 2.82) and 5–6 days (OR=1.59, 95% CI 0.93 to 2.73, p<0.1) compared with 0 days, participate in more out-of-school clubs (OR=1.06, 95% CI 1.02 to 1.10) and able to ride a bike (OR=1.39, 95% CI 1.00 to 1.93, reference: cannot ride a bike). Older children (OR=1.16, 95% CI 1.05 to 1.28) were more likely to test positive for SARS-CoV-2. Compared with boys, girls were more likely to test positive (OR=1.21, 95% CI 1.00 to 1.46), and compared with the most deprived (quintile 1), those living in the least deprived areas (quintile 4: OR=0.64, 95% CI 0.46 to 0.90; quintile 5: OR=0.64, 95% CI 0.46 to 0.89) were less likely to test positive for SARS-CoV-2. The model showed a low goodness of fit ($R^2$=0.02). Unadjusted multivariable logistic regression

**Table 2** Multivariable logistic regression model of significant health behaviour markers and probability of ever being PCR tested for SARS-CoV-2 between 1 March 2020 and 31 August 2021, accounting for baseline age, sex and deprivation, and clustered by school

| PCR tested for SARS-CoV-2 (n=6403, $R^2$=0.02) | OR | P value | 95% CI |
|---|---|---|---|
| Ate breakfast | 1.16* | 0.067 | 0.99 to 1.36 |
| *Reference: did not eat breakfast* | 1.00 | | |
| Actively travelled to school | 0.93 | 0.339 | 0.80 to 1.08 |
| *Reference: did not actively travel to school* | 1.00 | | |
| Actively travelled from school | 1.01 | 0.901 | 0.86 to 1.19 |
| *Reference: did not actively travel from school* | 1.00 | | |
| Number of fruit/vegetable portions | 1.00 | 0.959 | 0.97 to 1.03 |
| Number of times teeth brushed | 0.94 | 0.229 | 0.86 to 1.04 |
| Sleep hours | 1.01 | 0.682 | 0.97 to 1.04 |
| *Reference: 0 days physically active ≥60 min (previous 7 days)* | | | |
| 1–2 days physically active ≥60 min | 1.14 | 0.250 | 0.91 to 1.41 |
| 3–4 days physically active ≥60 min | 1.13 | 0.257 | 0.91 to 1.39 |
| 5–6 days physically active ≥60 min | 1.16 | 0.217 | 0.92 to 1.45 |
| 7 days physically active ≥60 min | 1.10 | 0.451 | 0.86 to 1.39 |
| *Reference: 0 days sedentary ≥2 hours (previous 7 days)* | 1.00 | | |
| 1–2 days sedentary ≥2 hours | 1.20 | 0.141 | 0.94 to 1.54 |
| 3–4 days sedentary ≥2 hours | 1.18 | 0.198 | 0.92 to 1.52 |
| 5–6 days sedentary ≥2 hours | 1.16 | 0.333 | 0.86 to 1.56 |
| 7 days sedentary ≥2 hours | 1.16 | 0.243 | 0.90 to 1.48 |
| *Reference: 0 days felt tired (previous 7 days)* | 1.00 | | |
| 1–2 days felt tired | 0.97 | 0.686 | 0.84 to 1.12 |
| 3–4 days felt tired | 1.00 | 0.963 | 0.85 to 1.16 |
| 5–6 days felt tired | 1.07 | 0.528 | 0.86 to 1.33 |
| 7 days felt tired | 0.97 | 0.728 | 0.83 to 1.14 |
| *Reference: 0 days consumed sugary snack (previous 7 days)* | 1.00 | | |
| 1–2 days consumed sugary snack | 1.24** | 0.018 | 1.04 to 1.49 |
| 3–4 days consumed sugary snack | 1.12 | 0.301 | 0.91 to 1.37 |
| 5–6 days consumed sugary snack | 1.31** | 0.008 | 1.07 to 1.61 |
| 7 days consumed sugary snack | 1.16 | 0.170 | 0.94 to 1.43 |
| Number of out-of-school club participations | 1.02* | 0.099 | 1.00 to 1.04 |
| Can ride a bike | 1.15* | 0.086 | 0.98 to 1.35 |
| *Reference: cannot ride a bike* | 1.00 | | |
| Can swim 25 m | 1.21** | 0.006 | 1.06 to 1.39 |
| *Reference: cannot swim 25 m* | 1.00 | | |
| Age on 1 March 2020 | 1.25** | < 0.001 | 1.16 to 1.35 |
| Sex (girl) | 0.92 | 0.161 | 0.81 to 1.04 |
| *Reference: sex (boy)* | 1.00 | | |
| *Reference: WIMD 2019 quintile 1 (Most deprived)* | 1.00 | | |
| WIMD 2019 quintile 2 | 0.95 | 0.600 | 0.80 to 1.14 |
| WIMD 2019 quintile 3 | 0.85* | 0.090 | 0.70 to 1.03 |
| WIMD 2019 quintile 4 | 0.87 | 0.131 | 0.73 to 1.04 |
| WIMD 2019 quintile 5 (Least deprived) | 0.85* | 0.078 | 0.72 to 1.02 |

*P<0.1; **p<0.05.
See online supplemental appendix 4 for variable codebook. Low to moderate correlation between variables (coefficients −0.19 to 0.71).
Complete case analysis.
WIMD, Welsh Index of Multiple Deprivation.

**Table 3** Multivariable logistic regression model of significant health behaviour markers and probability of ever PCR testing positive for SARS-CoV-2 between 1 March 2020 and 31 August 2021, accounting for baseline age, sex and deprivation, and clustered by school

| PCR test positive for SARS-CoV-2 (n=6403, $R^2$=0.02) | OR | P value | 95% CI |
|---|---|---|---|
| Ate breakfast<br>*Reference: did not eat breakfast* | 1.52**<br>1.00 | 0.043 | 1.01 to 2.27 |
| Actively travelled to school<br>*Reference: did not actively travel to school* | 0.91<br>1.00 | 0.481 | 0.70 to 1.18 |
| Actively travelled from school<br>*Reference: did not actively travel from school* | 0.98<br>1.00 | 0.910 | 0.72 to 1.33 |
| Number of fruit/vegetable portions | 0.98 | 0.461 | 0.94 to 1.03 |
| Number of times teeth brushed | 1.05 | 0.542 | 0.90 to 1.21 |
| Sleep hours | 0.97 | 0.345 | 0.92 to 1.03 |
| *Reference: 0 days physically active ≥60 min (previous 7 days)* | 1.00 | | |
| 1–2 days physically active ≥60 min | 1.69** | 0.035 | 1.04 to 2.74 |
| 3–4 days physically active ≥60 min | 1.76** | 0.018 | 1.10 to 2.82 |
| 5–6 days physically active ≥60 min | 1.59* | 0.091 | 0.93 to 2.73 |
| 7 days physically active ≥60 min | 1.50 | 0.158 | 0.85 to 2.65 |
| *Reference: 0 days sedentary ≥2 hours (previous 7 days)* | 1.00 | | |
| 1–2 days sedentary ≥2 hours | 0.96 | 0.847 | 0.63 to 1.47 |
| 3–4 days sedentary ≥2 hours | 0.94 | 0.789 | 0.59 to 1.50 |
| 5–6 days sedentary ≥2 hours | 0.93 | 0.803 | 0.51 to 1.68 |
| 7 days sedentary ≥2 hours | 1.02 | 0.946 | 0.63 to 1.65 |
| *Reference: 0 days felt tired (previous 7 days)* | 1.00 | | |
| 1–2 days felt tired | 1.18 | 0.207 | 0.91 to 1.51 |
| 3–4 days felt tired | 1.17 | 0.232 | 0.91 to 1.50 |
| 5–6 days felt tired | 1.19 | 0.243 | 0.89 to 1.60 |
| 7 days felt tired | 0.89 | 0.390 | 0.68 to 1.16 |
| *Reference: 0 days consumed sugary snack (previous 7 days)* | 1.00 | | |
| 1–2 days consumed sugary snack | 1.13 | 0.523 | 0.77 to 1.65 |
| 3–4 days consumed sugary snack | 1.06 | 0.783 | 0.70 to 1.61 |
| 5–6 days consumed sugary snack | 1.36 | 0.159 | 0.89 to 2.08 |
| 7 days consumed sugary snack | 1.08 | 0.727 | 0.71 to 1.63 |
| Number of out-of-school club participations | 1.06** | 0.002 | 1.02 to 1.10 |
| Can ride a bike<br>*Reference: cannot ride a bike* | 1.39**<br>1.00 | 0.049 | 1.00 to 1.93 |
| Can swim 25 m<br>*Reference: cannot swim 25 m* | 1.14 | 0.324 | 0.88 to 1.48 |
| Age on 1 March 2020 | 1.16** | 0.003 | 1.05 to 1.28 |
| Sex (girl)<br>*Reference: sex (boy)* | 1.21**<br>1.00 | 0.046 | 1.00 to 1.46 |
| Reference: WIMD 2019 quintile 1 (most deprived) | 1.00 | | |
| WIMD 2019 quintile 2 | 0.79 | 0.113 | 0.59 to 1.06 |
| WIMD 2019 quintile 3 | 0.79 | 0.128 | 0.59 to 1.07 |
| WIMD 2019 quintile 4 | 0.64** | 0.009 | 0.46 to 0.90 |
| WIMD 2019 quintile 5 | 0.64** | 0.008 | 0.46 to 0.89 |

*P<0.1; **p<0.05.
See online supplemental appendix 4 for variable codebook. Low to moderate correlation between variables (coefficients −0.19 to 0.71).
Complete case analysis.
WIMD, Welsh Index of Multiple Deprivation.

analyses are presented in online supplemental appendix 6.

## DISCUSSION

This study examines whether markers of health-related behaviours reported by primary school-age children between January 2018 and February 2020 are associated with the likelihood of ever being PCR tested for SARS-CoV-2 and ever testing positive between 1 March 2020 and 31 August 2021. Findings suggest that eating breakfast, weekly sugary snack consumption (both low and high), participating in more out-of-school clubs, being able to ride a bike and being able to swim 25 m were associated with an increased likelihood of being tested for SARS-CoV-2. Health behaviours associated with an increased likelihood of testing positive for SARS-CoV-2 were eating breakfast, engaging in higher weekly physical activity, participating in more out-of-school clubs and riding a bike. Boys were more likely to test positive for SARS-CoV-2 than girls, and those living in a less deprived area were less likely to test positive than those residing in the most deprived area.

This study encompasses a period of both targeted and mass PCR testing, and detecting positive child cases using routine PCR testing data in this study requires a parent/guardian to take the child for testing. We find associations between child-reported health-related behaviours with both PCR testing for SARS-CoV-2 and testing positive for SARS-CoV-2. Through this, we theorise that because health behaviours are largely guided and facilitated by parents, our associations are likely to be reflecting health literacy among parents, along with monitoring behaviours. In the case of symptomatic testing, the detection of positive child cases relies on parents recognising symptoms and communication with their child. For asymptomatic testing through the use of LFT (eg, asymptomatic school testing between February and April 2021), guidance encouraged positive LFTs to be followed up with PCR testing, requiring knowledge of how to access testing services and ability to access services (eg, transport). These behaviours form a level of health literacy, recognised as the ability to access, understand, interpret and apply medical information and make informed decisions regarding medical advice, issues or guidelines.[39] Parental health literacy impacts the decision a parent makes relating to their child[40] and is correlated with a number of health indicators including knowledge of health and health services, and the parent and child engaging in health-promoting behaviours.[8 39]

Parenting is an important contributor to promoting positive health behaviours in children, and is represented by a constellation of attitudes, behaviours and values for the child. The presence of multiple physically active behaviours represented by the association of being able to swim, ride a bike and participation in more out-of-school clubs may represent underlying parental involvement and modelling behaviour, including involvement in leisure-time activities, providing financial and transport provision to attend organised activities such as access to swimming lessons and the provision of equipment.[7] This may also have a socioeconomic component, building on the ideas of Bourdieu in terms of social capital, and accessing health-enhancing material items.[41]

Diet-related findings of eating breakfast and restrictive weekly sugary snack consumption (1–2 days/week) may indicate higher parental monitoring, supporting our theory. In comparison, higher weekly sugary snack consumption (5–6 days/week) may represent less restrictive parental monitoring and more autonomy and choice for the child. We posit that as parental behaviours are often driven by underlying styles of parenting,[42] the associations could be depicting varying levels of control; for instance, those snacking one to two times perhaps have parents with greater control versus those snacking five to six times with parents with less controlling styles. This theory may well transcend into other behaviours, including limits and freedom in socialising with others, placing a greater likelihood of infection of illness—including COVID-19.

While evidence recognises the importance of adequate nutrition[16] and physical activity[17] for cardiometabolic health and immune system function, the findings in the current study draw attention to another potential mechanism of increased contacts and exposure to SARS-CoV-2. Engagement in physically active behaviours such as out-of-school clubs, higher frequency of physically active days in a week and riding a bike may increase the number of social contacts of the child. Indeed, there is a wealth of evidence demonstrating that childhood physical activity participation is highly influenced by their social environment and coparticipation with peers.[43] It is therefore possible that physically active children had increased social contacts and exposure to SARS-CoV-2 through coparticipation of activity and play opportunities.

However, it is important to note that physical activity is an essential health behaviour required for optimal development and a range of health and well-being outcomes. These findings must be considered in balance with the importance of encouraging these behaviours and providing physically active opportunities during childhood. This viewpoint was also reflected in government guidance and risk assessments during the COVID-19 pandemic through the reopening of children's playgrounds and outdoor play spaces, with explicit reference to outdoor play and physical activity as fundamental for children's development and well-being.[44]

Contact patterns may also explain sex differences observed in this study, as we found girls are more likely to test positive for SARS-CoV-2. In addition to age assortative mixing patterns of children, there is a developmental tendency by children to socially interact with members of the same sex and engage in gender-type activity.[45] For girls, the location of play preferences is more likely to be indoors and in contact with supervising adults, where exposure to SARS-CoV-2 is possibly greater.[46] The findings of association between increasing age and likelihood

of testing positive for SARS-CoV-2 in this study are supported by wider literature which suggests increasing susceptibility of infection in the adolescent age group compared with younger than 10–14 years.[47]

Our findings also show an area-level social gradient. Those living in the least deprived areas (WIMD quintiles 4 and 5) were less likely to test positive for SARS-CoV-2 compared with the most deprived (quintile 1), which may reflect deprivation-related exposure patterns to SARS-CoV-2. Indeed, research conducted using the WIMD and English area-level deprivation indicators found adults living in the most deprived areas demonstrated differential exposures to SARS-CoV-2.[48] This included patterns of public activities such as attending work or education outside of the household, using public transport and car sharing with non-household members. This, and considerations of the deprivation-related disparities in the built environment including access to open spaces, highlights the inequalities that persist in SARS-CoV-2 infection. Furthermore, while it is likely that children mix with others from similar demographic areas, the finding in our study may also reflect community prevalence which was not captured.

## CONCLUSION

We theorise that health-promoting behaviours associated with a child being tested for SARS-CoV-2 and being identified as positive may be a proxy of higher parental health literacy and monitoring behaviours. Furthermore, coparticipation in physically active behaviours with peers may increase exposure to SARS-CoV-2. This must be considered from a risk-versus-benefit approach in relation to promoting these health behaviours, given the importance of health-related behaviours such as physical activity during childhood for development and wellbeing. This national-level case study using survey data linked with routine health data in Wales provides insight into these issues from a devolved policy-making context, with the potential for replicability and portability to other jurisdictions.

**Author affiliations**
[1]Population Data Science, Medical School, Swansea University, Swansea, UK
[2]Department of Education and Childhood Studies, Swansea University, Swansea, UK
[3]National Centre for Population Health and Wellbeing Research, Medical School, Swansea University, Swansea, UK
[4]Department of Nursing and Health Science, Fulda University of Applied Sciences, Fulda, Germany
[5]Center for Applied Health Sciences, Leuphana University of Lüneburg, Lüneburg, Germany
[6]Faculty of Sport and Health Sciences, Technical University of Munich, München, Germany

**Twitter** Emily Marchant @emily_marchant, Tom Crick @ProfTomCrick, Lucy J Griffiths @LJ_Griffiths, Richard Fry @richfry, Orkan Okan @orkanokan_, Ashley Akbari @AshleyAkbari and Sinead Brophy @@SineadBr

**Acknowledgements** The authors would like to thank all the participating primary schools and pupils who took part in this study. This work was supported by the National Centre for Population Health and Wellbeing Research through the *HAPPEN* network. This study makes use of anonymised data held in the Secure Anonymised Information Linkage (SAIL) Databank. We would like to acknowledge all the data providers who make anonymised data available for research. We would also like to thank Dr Annemarie Docherty and Dr Olivia Swann from The University of Edinburgh for providing informal peer review input to the final draft.

**Contributors** EM and SB conceptualised the study design. EM and MJ acquired the data, and EM and JK were responsible for data curation. EM performed the statistical analysis, undertook the initial interpretation of the data and wrote the initial draft. EL and SB contributed to the writing of the manuscript and provided statistical guidance. EL, JK, SB, LC and RL provided critical interpretation of the data. The manuscript was critically reviewed and edited by EL, TC, LJG, RF, KD, OO, MJ, LC, FT, JK, AA, RL and SB. SB provided supervision and TC and LJG provided mentorship. EM is the guarantor of the study. EM, EL, TC, LJG, RF, KD, OO, MJ, LC, FT, JK, AA, RL and SB approved the final manuscript and agreed to be accountable for all aspects of the work.

**Funding** The Economic and Social Research Council (ESRC) funded the development of the *HAPPEN* network (grant number: ES/J500197/1) which this research was conducted through. The National Centre for Population Health and Wellbeing Research (NCPHWR) funded by Health and Care Research Wales provided infrastructural support for this work. This work was supported by the Con-COV team funded by the Medical Research Council (grant number: MR/V028367/1). This work was supported by Health Data Research UK, which receives its funding from HDR UK (HDR-9006) funded by the UK Medical Research Council, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Department of Health and Social Care (England), Chief Scientist Office of the Scottish Government Health and Social Care Directorates, Health and Social Care Research and Development Division (Welsh Government), Public Health Agency (Northern Ireland), British Heart Foundation (BHF) and the Wellcome Trust. This work was a collaboration with the ADR Wales programme of work. ADR Wales is part of the Economic and Social Research Council (part of UK Research and Innovation) funded ADR UK (grant number: ES/S007393/1). This work was supported by the Wales COVID-19 Evidence Centre, funded by Health and Care Research Wales and by the COVID-19 Longitudinal Health and Wellbeing National Core Study funded by the Medical Research Council (MC_PC_20030).

**Competing interests** None declared.

**Patient and public involvement** Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

**Patient consent for publication** Not applicable.

**Ethics approval** This study involves human participants and was approved by the Swansea University Medical School Research Ethics Committee (reference ID: 2017-0033H). Participants gave informed consent to participate in the study before taking part.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available upon reasonable request. The routine data used in this study are available in the SAIL Databank at Swansea University, Swansea, UK. All proposals to use SAIL data are subject to review by an IGRP. Before any data can be accessed, approval must be given by the IGRP. The IGRP gives careful consideration to each project to ensure proper and appropriate use of SAIL data. When access has been approved, it is gained through a privacy-protecting safe haven and remote access system referred to as the SAIL Gateway. SAIL has established an application process to be followed by anyone who would like to access data via SAIL: https://www.saildatabank.com/application-process. This study has been approved by the IGRP as project 0911.

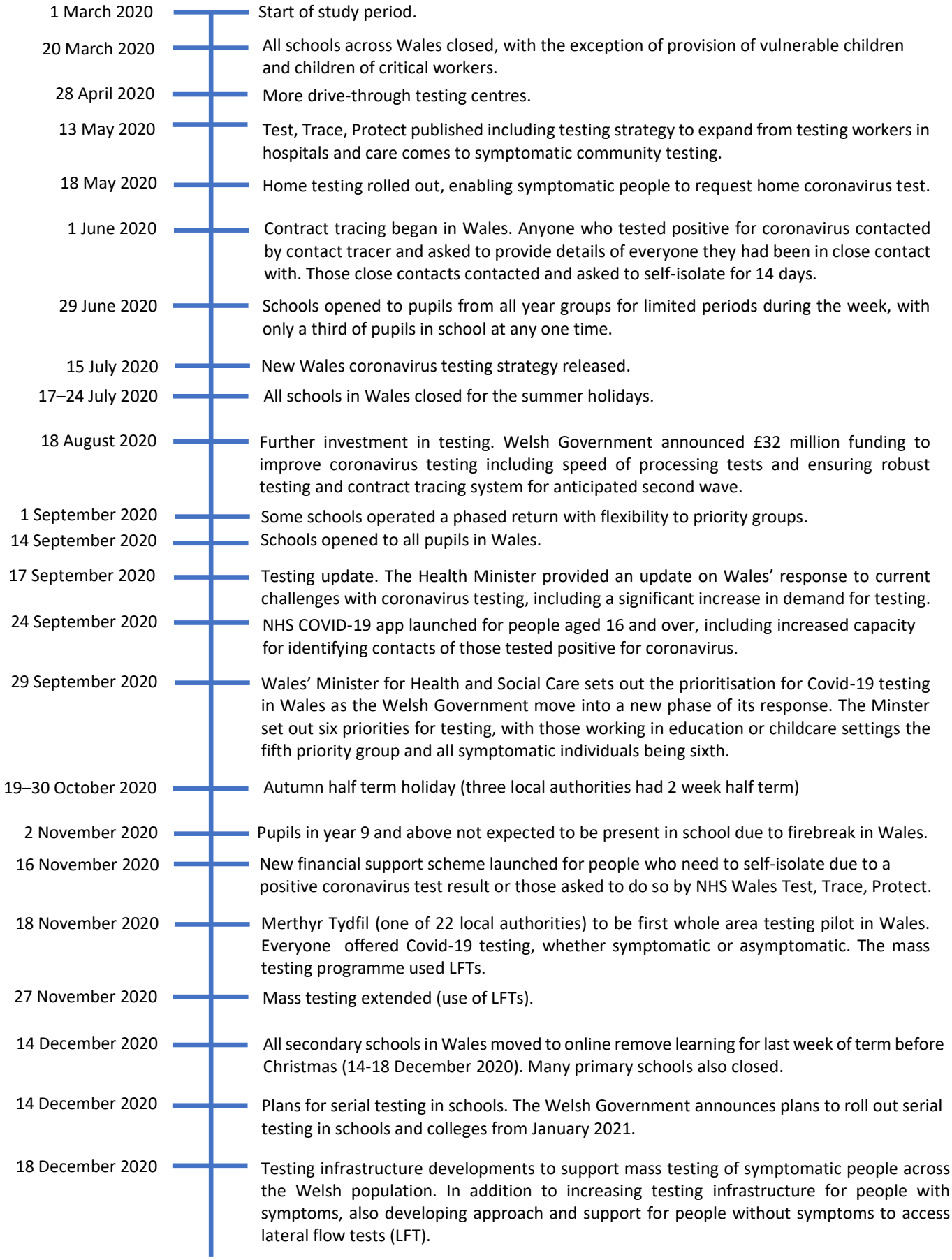
**ORCID iDs**
Emily Marchant http://orcid.org/0000-0002-9701-5991
Emily Lowthian http://orcid.org/0000-0001-9362-0046
Tom Crick http://orcid.org/0000-0001-5196-9389
Lucy J Griffiths http://orcid.org/0000-0001-9230-624X
Richard Fry http://orcid.org/0000-0002-7968-6679
Michaela James http://orcid.org/0000-0001-7047-0049
Jonathan Kennedy http://orcid.org/0000-0002-1122-6502
Ashley Akbari http://orcid.org/0000-0003-0814-0801
Ronan Lyons http://orcid.org/0000-0001-5225-000X
Sinead Brophy http://orcid.org/0000-0001-7417-2858

## REFERENCES

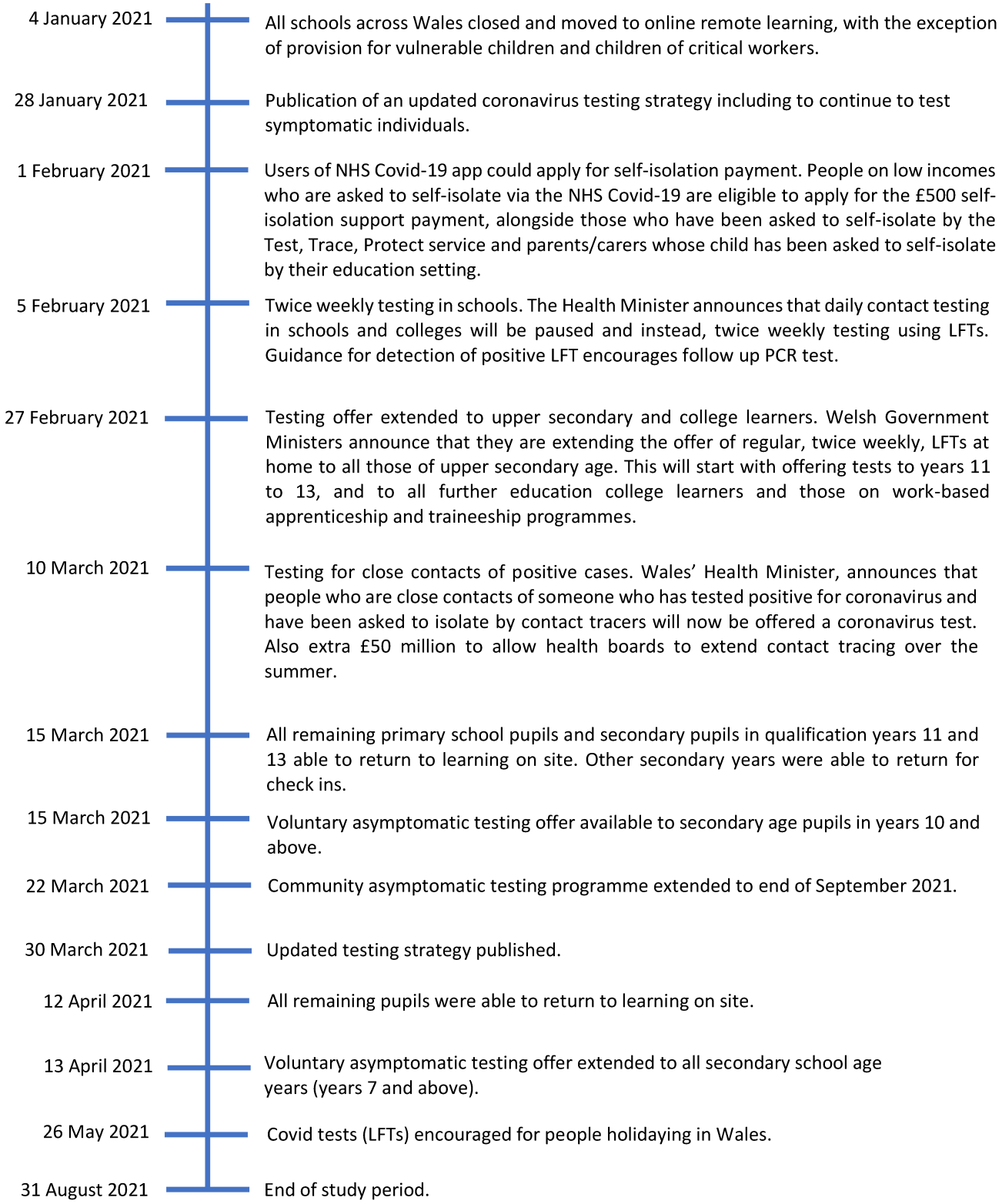
1 James M, Marchant E, Defeyter MA, *et al*. Impact of school closures on the health and well-being of primary school children in Wales UK: a routine data linkage study using the happen survey (2018–2020). *BMJ Open* 2021;11:e051574.
2 Marchant E, Todd C, James M, *et al*. Primary school staff perspectives of school closures due to COVID-19, experiences of schools reopening and recommendations for the future: a qualitative survey in Wales. *PLoS One* 2021;16:e0260396.
3 Craigie AM, Lake AA, Kelly SA, *et al*. Tracking of obesity-related behaviours from childhood to adulthood: a systematic review. *Maturitas* 2011;70:266–84.
4 OECD. Are children active and physically healthy? In: *Measuring what matters for child well-being and policies*. Paris: OECD Publishing, 2021: 1–295.
5 OECD. Education in the Digital Age. In: Burns T, Gottschalk F, eds. *Education in the digital age*. OECD, 2020: 1–218.
6 Pyper E, Harrington D, Manson H. The impact of different types of parental support behaviours on child physical activity, healthy eating, and screen time: a cross-sectional study. *BMC Public Health* 2016;16:568.
7 Edwardson CL, Gorely T. Parental influences on different types and intensities of physical activity in youth: a systematic review. *Psychol Sport Exerc* 2010;11:522–35.
8 Pawellek M, Kopf FM, Egger N. Pathways linking parental health literacy with health behaviours directed at the child: a scoping review. *Health Promot Int* 2021;2021:1–29.
9 Moore SA, Faulkner G, Rhodes RE, *et al*. Impact of the COVID-19 virus outbreak on movement and play behaviours of Canadian children and youth: a national survey. *Int J Behav Nutr Phys Act* 2020;17.
10 Chowdhury R, van Daalen KR, Franco OH. Cardiometabolic health: key in reducing adverse COVID-19 outcomes. *Glob Heart* 2020;15:58.
11 Koch R, Jordan S, Starker A. Health behaviour and COVID-19: initial findings on the pandemic. *J Heal Monit* 2020;2020:1–14.
12 Childs CE, Calder PC, Miles EA. Diet and immune function. *Nutrients* 2019;11:1933.
13 Lange KW, Nakamura Y. Lifestyle factors in the prevention of COVID-19. *Glob Health J* 2020;4:146–52.
14 Chastin SFM, Abaraogu U, Bourgois JG, *et al*. Effects of regular physical activity on the immune system, vaccination and risk of community-acquired infectious disease in the general population: systematic review and meta-analysis. *Sports Med* 2021;51:1673–86.
15 Furman D, Campisi J, Verdin E, *et al*. Chronic inflammation in the etiology of disease across the life span. *Nat Med* 2019;25:1822–32.
16 Wintergerst ES, Maggini S, Hornig DH. Contribution of selected vitamins and trace elements to immune function. *Ann Nutr Metab* 2007;51:301–23.
17 Laddu DR, Lavie CJ, Phillips SA, *et al*. Physical activity for immunity protection: Inoculating populations with healthy living medicine in preparation for the next pandemic. *Prog Cardiovasc Dis* 2021;64:102–4.
18 Hamer M, Kivimäki M, Gale CR, *et al*. Lifestyle risk factors, inflammatory mechanisms, and COVID-19 hospitalization: a community-based cohort study of 387,109 adults in UK. *Brain Behav Immun* 2020;87:184–7.
19 Lee SW, Lee J, Moon SY, *et al*. Physical activity and the risk of SARS-CoV-2 infection, severe COVID-19 illness and COVID-19 related mortality in South Korea: a nationwide cohort study. *Br J Sports Med* 2022;56:18.
20 Zhou B, Yuan Y, Wang S, *et al*. Risk profiles of severe illness in children with COVID-19: a meta-analysis of individual patients. *Pediatr Res* 2021;90:347–52.
21 Shekerdemian LS, Mahmood NR, Wolfe KK, *et al*. Characteristics and outcomes of children with coronavirus disease 2019 (COVID-19) infection admitted to US and Canadian pediatric intensive care units. *JAMA Pediatr* 2020;174:868–73.
22 Castagnoli R, Votto M, Licari A, *et al*. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection in children and adolescents. *JAMA Pediatr* 2020;174:882–9.
23 OECD. Combatting COVID-19's effect on children - OECD, 2020. Available: https://www.oecd.org/coronavirus/policy-responses/combatting-covid-19-s-effect-on-children-2e1f3b2f/ [Accessed 01 Oct 2021].
24 Senedd Research. Coronavirus timeline: Welsh and UK governments' response – IN BRIEF, 2020. Available: https://seneddresearch.blog/2020/03/19/covid-19-timeline-welsh-and-uk-governments-response/ [Accessed 30 Mar 2022].
25 Todd C, Christian D, Tyler R, *et al*. Health and attainment of pupils involved in a primary education network): working in partnership to improve child health and education. *Perspect Public Health* 2016;136:115–6.
26 Christian D, Todd C, Davies H, *et al*. Community led active schools programme (CLASP) exploring the implementation of health interventions in primary schools: headteachers' perspectives. *BMC Public Health* 2015;15:238.
27 Todd C, Christian D, Davies H. Headteachers' prior beliefs on child health and their engagement in school based health interventions: a qualitative study. *BMC Res Notes* 2015;18:161.
28 Marchant E, Todd C, Cooksey R, *et al*. Curriculum-based outdoor learning for children aged 9-11: A qualitative analysis of pupils' and teachers' views. *PLoS One* 2019;14:e0212242.
29 Marchant E, Todd C, Stratton G. The daily mile: Whole-school recommendations for implementation and sustainability. *A mixed-methods study. PLoS One* 2020;15.
30 Marchant E, Griffiths L, Crick T, *et al*. COVID-19 mitigation measures in primary schools and association with infection and school staff wellbeing: an observational survey linked with routine data in Wales, UK. *PLoS One* 2022;17:e0264023.
31 Everson B, Mackintosh KA, McNarry MA, *et al*. Can Wearable Cameras be Used to Validate School-Aged Children's Lifestyle Behaviours? *Children* 2019;6:20.
32 Department of Health & Social Care. COVID-19 testing data: methodology note - GOV.UK, 2020. Available: https://www.gov.uk/government/publications/coronavirus-covid-19-testing-data-methodology/covid-19-testing-data-methodology-note [Accessed 30 Mar 2022].
33 Ford DV, Jones KH, Verplancke J-P, *et al*. The Sail Databank: building a national architecture for e-health research and evaluation. *BMC Health Serv Res* 2009;9:157.
34 Jones KH, Ford DV, Thompson S, *et al*. A profile of the Sail Databank on the UK secure research platform. *Int J Popul Data Sci* 2019;4:1134.
35 Lyons J, Akbari A, Torabi F, *et al*. Understanding and responding to COVID-19 in Wales: protocol for a privacy-protecting data platform for enhanced epidemiology and evaluation of interventions. *BMJ Open* 2020;10:43010.
36 Benchimol EI, Smeeth L, Guttmann A, *et al*. The reporting of studies conducted using observational Routinely-collected health data (record) statement. *PLoS Med* 2015;12:e1001885.
37 Welsh Government. Welsh index of multiple deprivation, 2019. Available: https://gov.wales/welsh-index-multiple-deprivation-full-index-update-ranks-2019
38 Dong Y, Peng C-YJ. Principled missing data methods for researchers. *Springerplus* 2013;2:1–17.
39 Sørensen K, Van den Broucke S, Fullam J, *et al*. Health literacy and public health: a systematic review and integration of definitions and models. *BMC Public Health* 2012;12.
40 Morrison AK, Glick A, Yin HS. Health literacy: implications for child health. *Pediatr Rev* 2019;40:263–77.
41 Bourdieu P. The Forms of Capital. In: *Handbook of theory and research for the sociology of education*. New York: Greenwood, 1986: 241–58.
42 Darling N, Steinberg L. Parenting style as context: an integrative model. *Psychol Bull* 1993;113:487–96.
43 Li L, Moosbrugger ME. Correlations between physical activity participation and the environment in children and adolescents: a systematic review and meta-analysis using ecological frameworks. *Int J Environ Res Public Health* 2021;18:9080:.
44 Welsh Government. Reopening children's playgrounds and outdoor play areas: coronavirus. Guidance for those in charge of playgrounds or outdoor play areas to keep children safe during the coronavirus pandemic, 2020. Available: https://gov.wales/reopening-childrens-playgrounds-and-outdoor-play-areas-coronavirus-html [Accessed 25 Apr 2022].

45 Mehta CM, Strough J. Sex segregation in friendships and normative contexts across the life span. *Developmental Review* 2009;29:201–20.
46 Edwards CP, Knoche L, Kumru A. Play patterns and gender. *Encycl women Gend Sex similarities Differ impact Soc Gend* 2001:809–15 http://digitalcommons.unl.edu/psychfacpub/607/
47 Viner RM, Mytton OT, Bonell C, *et al*. Susceptibility to SARS-CoV-2 infection among children and adolescents compared with adults: a systematic review and meta-analysis. *JAMA Pediatr* 2021;175:143-156.
48 Beale S, Braithwaite I, Navaratnam AMD, *et al*. Deprivation and exposure to public activities during the COVID-19 pandemic in England and Wales. *J Epidemiol Community Health* 2022;76:319–26.

Online supplemental appendix 1: see [24]



Online supplemental appendix 1: see [24]



**Online supplemental appendix 2: RECORD statement**

**The RECORD statement – checklist of items, extended from the STROBE statement, that should be reported in observational studies using routinely collected health data.**

| | Item No. | STROBE items | Location in manuscript where items are reported | RECORD items | Location in manuscript where items are reported |
|---|---|---|---|---|---|
| **Title and abstract** | | | | | |
| | 1 | (a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found | | RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included.<br><br>RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract.<br><br>RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract. | 1.1: Title and abstract<br><br>1.2: Title and abstract |
| **Introduction** | | | | | |
| Background rationale | 2 | Explain the scientific background and rationale for the investigation being reported | | | Background |
| Objectives | 3 | State specific objectives, including any prespecified hypotheses | | | Background |
| **Methods** | | | | | |
| Study Design | 4 | Present key elements of study design early in the paper | | | Methods - Study design |
| Setting | 5 | Describe the setting, locations, and relevant dates, including | | | Methods - Study design |

**Online supplemental appendix 2: RECORD statement**

| | | | | | |
|---|---|---|---|---|---|
| | | periods of recruitment, exposure, follow-up, and data collection | | | |
| Participants | 6 | *(a) Cohort study* - Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up<br>*Case-control study* - Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls<br>*Cross-sectional study* - Give the eligibility criteria, and the sources and methods of selection of participants<br><br>*(b) Cohort study* - For matched studies, give matching criteria and number of exposed and unexposed<br>*Case-control study* - For matched studies, give matching criteria and the number of controls per case | | RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided.<br><br>RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided.<br><br>RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage. | 6.1: Figure 1: Cohort Flow Diagram<br><br>6.3: Figure 1: Cohort Flow Diagram |
| Variables | 7 | Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable. | | RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided. | 7.1: Supplemental appendix 4: HAPPEN survey variable codebook |
| Data sources/ measurement | 8 | For each variable of interest, give sources of data and details of methods of assessment (measurement). | | | Methods - The HAPPEN survey and linked SAIL data |

**Online supplemental appendix 2: RECORD statement**

| | | | | | |
|---|---|---|---|---|---|
| | | Describe comparability of assessment methods if there is more than one group | | | |
| Bias | 9 | Describe any efforts to address potential sources of bias | | | Methods - Quantitative analysis |
| Study size | 10 | Explain how the study size was arrived at | | | Figure 1: Cohort flow diagram |
| Quantitative variables | 11 | Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why | | | Methods - Quantitative analysis |
| Statistical methods | 12 | (a) Describe all statistical methods, including those used to control for confounding<br>(b) Describe any methods used to examine subgroups and interactions<br>(c) Explain how missing data were addressed<br>(d) *Cohort study* - If applicable, explain how loss to follow-up was addressed<br>*Case-control study* - If applicable, explain how matching of cases and controls was addressed<br>*Cross-sectional study* - If applicable, describe analytical methods taking account of sampling strategy<br>(e) Describe any sensitivity analyses | | | Methods - Quantitative analysis |
| Data access and cleaning methods | | .. | | RECORD 12.1: Authors should describe the extent to which the | 12.1: Methods - The HAPPEN |

**Online supplemental appendix 2: RECORD statement**

| | | | | | |
|---|---|---|---|---|---|
| | | | | investigators had access to the database population used to create the study population.<br><br>RECORD 12.2: Authors should provide information on the data cleaning methods used in the study. | survey and linked SAIL data<br><br>12.2: Figure 1 – Cohort flow diagram |
| Linkage | | .. | | RECORD 12.3: State whether the study included person-level, institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided. | 12.3: Methods - Study design<br><br>The HAPPEN survey and linked SAIL data |
| **Results** | | | | | |
| Participants | 13 | (a) Report the numbers of individuals at each stage of the study (*e.g.*, numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed)<br>(b) Give reasons for non-participation at each stage.<br>(c) Consider use of a flow diagram | | RECORD 13.1: Describe in detail the selection of the persons included in the study (*i.e.,* study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram. | 13.1: Methods - Quantitative analysis<br><br>Figure 1: Cohort flow diagram |
| Descriptive data | 14 | (a) Give characteristics of study participants (*e.g.*, demographic, clinical, social) and information on exposures and potential confounders<br>(b) Indicate the number of participants with missing data for each variable of interest<br>(c) *Cohort study* - summarise follow-up time (*e.g.*, average and total amount) | | | Results - Table 1 Descriptive statistics<br>Full descriptive statistics table: Online supplemental appendix 5 |

**Online supplemental appendix 2: RECORD statement**

| | | | | | |
|---|---|---|---|---|---|
| Outcome data | 15 | *Cohort study* - Report numbers of outcome events or summary measures over time<br>*Case-control study* - Report numbers in each exposure category, or summary measures of exposure<br>*Cross-sectional study* - Report numbers of outcome events or summary measures | | | Results - Table 1 Descriptive statistics<br>Full descriptive statistics table: Online supplemental appendix 6 |
| Main results | 16 | (a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included<br>(b) Report category boundaries when continuous variables were categorized<br>(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period | | | Results – Table 3 and table 4<br><br>Online supplemental appendix 6: Unadjusted multivariable logistic regression analyses |
| Other analyses | 17 | Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses | | | Results |
| **Discussion** | | | | | |
| Key results | 18 | Summarise key results with reference to study objectives | | | Results |
| Limitations | 19 | Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias | | RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, | Strengths and limitations |

**Online supplemental appendix 2: RECORD statement**

| | | | | | |
|---|---|---|---|---|---|
| | | | | unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported. | |
| Interpretation | 20 | Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence | | | Discussion Conclusion |
| Generalisability | 21 | Discuss the generalisability (external validity) of the study results | | | Discussion Conclusion Strengths and limitations |
| **Other Information** | | | | | |
| Funding | 22 | Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based | | | Funding |
| Accessibility of protocol, raw data, and programming code | | .. | | RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code. | Availability for data and materials |

*Reference: Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM, the RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Medicine* 2015; in press.

# THE HAPPEN SURVEY

* Required

Consent Form

Before you start please click this link to read the information sheet -> https://happen-wales.co.uk/wp-content/uploads/2019/02/Child-Consent-2019.pdf

1. I have read the child information sheet -> https://happen-wales.co.uk/wp-content/uploads/2019/02/Child-Consent-2019.pdf (click the link if you haven't read it) and understand that if I take part I can change my mind at any time, and this will not be a problem at all. *

*Mark only one oval.*

- ◯ Yes
- ◯ No

2. I am happy for you to use my questionnaire for research. Only the researchers in the team will know my name and will not tell anyone else my answers *

*Mark only one oval.*

- ◯ Yes
- ◯ No do not use my questionnaire

3. I am happy for you to look at my school and health records to see how my school is doing (as a group). This is anonymous which means I cannot be identified *

*Mark only one oval.*

- ◯ Yes
- ◯ No

If you do not wish to take part in the questionnaire please do not continue.

Please click next to start the questionnaire!

## About You

4. First Name *

5. Last Name *

6. Home Post Code *

7. What school do you go to? *

8. What year are you in? *

*Mark only one oval.*

- ◯ Year 4
- ◯ Year 5
- ◯ Year 6
- ◯ Year 7

9. Gender *

*Mark only one oval.*

- ◯ Boy
- ◯ Girl
- ◯ Prefer not to say

Date of Birth

10. Year *

*Mark only one oval.*

- ◯ 2007
- ◯ 2008
- ◯ 2009
- ◯ 2010
- ◯ 2011
- ◯ 2012

11. Month *

*Mark only one oval.*

- January
- February
- March
- April
- May
- June
- July
- August
- September
- October
- November
- December

12. Day *

*Mark only one oval.*

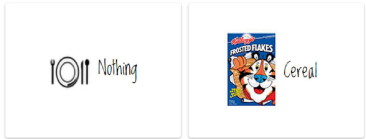
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20
- 21
- 22
- 23
- 24
- 25
- 26
- 27
- 28
- 29
- 30
- 31

YESTERDAY


Firstly, think carefully about what you did YESTERDAY and then answer the following questions....

13. 1. What did you eat for breakfast YESTERDAY? *

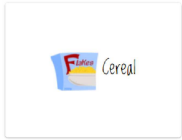
*Check all that apply.*




Nothing




Sugary cereal e.g. cocopops, frosties, sugar puffs, chocolate cereals

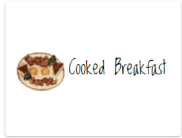



Healthy cereal e.g. porridge, weetabix, readybrek, muesli, branflakes, cornflakes


Snacks


Fruit


Toast




Cooked breakfast

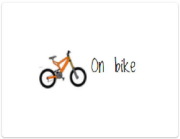

Yoghurt

Other:

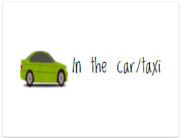
14. 2. How did you get to school YESTERDAY morning? *

*Mark only one oval.*


On the bus




On bike




In the car/taxi


Walked


Ran/jogged


Scooter


Skateboarded/Rollerbladed


15. 3. What did you have to eat for lunch YESTERDAY? *

*Mark only one oval.*

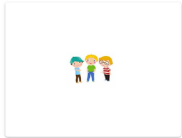
School dinner
Packed lunch
Nothing

16. 4. What did you do for MOST of your break-times YESTERDAY? (This includes lunchtime) *

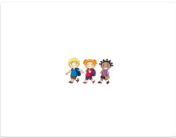
*Mark only one oval.*


Sat around inside or outside


Ran around



Stood around



Walked around


17. 5. Do you have an afternoon break at school? *

*Mark only one oval.*

YES
NO

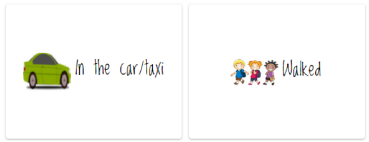
18. 6. How did you get home YESTERDAY? *

*Mark only one oval.*


On the bus




On bike




In the car/taxi

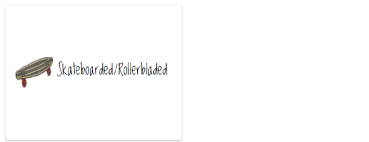



Walked


Ran/jogged




Scooter




Skateboarded/Rollerbladed

AFTER SCHOOL

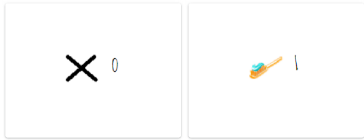

19. 7. How many portions of fruit and vegetables did you eat YESTERDAY? *


*Mark only one oval.*

- 0
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8

20. 8. How many times did you brush your teeth YESTERDAY? *

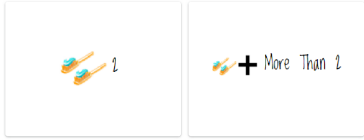
*Mark only one oval.*




0




1




2




3

21. 9. What time did you fall asleep YESTERDAY (to the nearest half hour)? *

*Mark only one oval.*

- 7:00pm
- 7:30pm
- 8:00pm
- 8:30pm
- 9:00pm
- 9:30pm
- 10:00pm
- 10:30pm
- 11:00pm
- 11:30pm
- 12:00am
- 12:30am
- 1:00am
- 1:30am
- 2:00am
- 3:00am
- 3:30am
- 4:00am

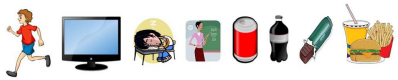
22. 10. What time did you wake up TODAY (to the nearest half hour)? *

*Mark only one oval.*

- 5:00am
- 5:30am
- 6:00am
- 6:30am
- 7:00am
- 7:30am
- 8:00am
- 8:30am
- 9:00am

## THE LAST WEEK

**NOW think about what you did in the last 7 days...**



23. 11a. In the last 7 days, how many days did you do sports or exercise for at least 1 hour in total (This includes doing any activities or playing sports where your heart beat faster, you breathed faster and you felt warmer? *

*Mark only one oval.*

- 0 days
- 1-2 days
- 3-4 days
- 5-6 days
- 7 days

24. 11b. In the last 7 days, how many days did you watch TV/play online games/use the internet etc. for 2 or more hours a day (in total)? *

*Mark only one oval.*

- 0 days
- 1-2 days
- 3-4 days
- 5-6 days
- 7 days

25. 11c. In the last 7 days, how many days did you feel tired? *

*Mark only one oval.*

- 0 days
- 1-2 days
- 3-4 days
- 5-6 days
- 7 days

26. 11d. In the last 7 days, how many days did you feel like you could concentrate/pay attention well in class? *

*Mark only one oval.*

- 0 days
- 1-2 days
- 3-4 days
- 5-6 days
- 7 days

27. 11e. In the last 7 days, how many days did you drink at least one fizzy drink (e.g. coke, fanta, sprite) *

*Mark only one oval.*

- 0 days
- 1-2 days
- 3-4 days
- 5-6 days
- 7 days

28. 11f. In the last 7 days, how many days did you eat at least one sugary snack (e.g. chocolate bar, sweets) *

*Mark only one oval.*

- 0 days
- 1-2 days
- 3-4 days
- 5-6 days
- 7 days

29. 11g. In the last 7 days, how many days did you eat take away foods (e.g. McDonalds, KFC, chinese) *

*Mark only one oval.*

- 0 days
- 1-2 days
- 3-4 days
- 5-6 days
- 7 days

## Sport and Activity

30. 12. These questions are going to ask you how you feel about physical activity (This includes any activity where your heart beats faster, you breathe faster and you feel warmer) *


*Mark only one oval per row.*

| | Strongly agree | Agree | Disagree | Strongly disagree |
|---|---|---|---|---|
| I want to take part in physical activity | ○ | ○ | ○ | ○ |
| I feel confident to take part in lots of different physical activities | ○ | ○ | ○ | ○ |
| I am good at lots of different physical activities | ○ | ○ | ○ | ○ |
| I understand why taking part in physical activity is good for me | ○ | ○ | ○ | ○ |

31. 13a. How many times do you take part in a sports club OUTSIDE OF SCHOOL each week?

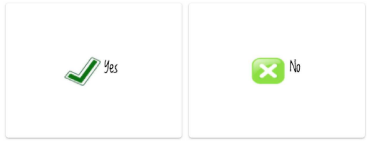
*Mark only one oval.*

0 1 2 3 4 5 6 7 8 9 10

32. 13b. If you take part in a sports club OUTSIDE of school, what is the name of the sports club? (For example Swansea Rugby Club Under 11's)

33. 14. Are you a member of cubs, brownies, scouts or guides? *

*Mark only one oval.*




Yes

No

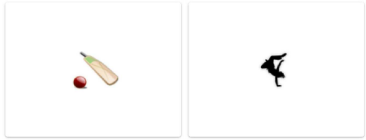
34. 15. Which of these sports or physical activities would you MOST like to try? (That you haven't tried before) *

*Mark only one oval.*

Athletics



Basketball



Cricket

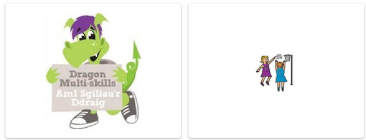


Dance

Gymnastics



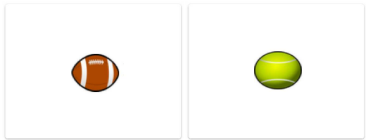
Hockey




Multi Skills



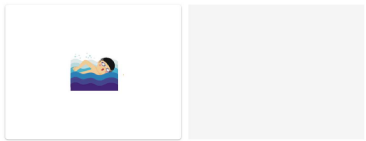
Netball



Rugby



Tennis



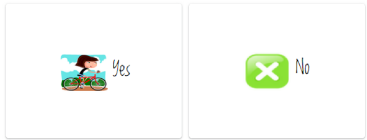
Swimming

I do not want to try anything-I don't like sport or activity

Other: _____

35. 16. Can you ride a bike WITHOUT STABILISERS? *

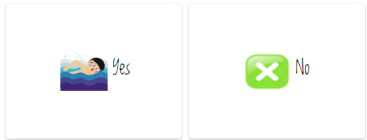
*Mark only one oval.*




Yes No

36. 17. Can you swim 25 metres WITHOUT A FLOAT OR ARMBANDS? (This is 1 length of a standard swimming pool) *

*Mark only one oval.*




Yes No

## You and your feelings

This part of the survey is going to ask you how you feel. There are no right or wrong answers. You should just pick the answer which is best for you.

37. 18. Tell us if you agree or disagree with the following: *


*Mark only one oval per row.*

| | Strongly agree | Agree | Don't agree or disagree | Disagree | Strongly disagree |
|---|---|---|---|---|---|
| I am doing well at school | | | | | |
| I have lots of choice over things that are important to me | | | | | |
| There are lots of things I'm good at | | | | | |

19. On a scale of 0 to 10 (0 being very unhappy and 10 being very happy, how do you feel about:
*Based on the Good Childhood Index by the Children's Society

38. Your Health *

*Mark only one oval.*

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Very unhappy | | | | | | | | | | | | Very happy |

39. Your School *

*Mark only one oval.*

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Very unhappy | | | | | | | | | | | | Very happy |

40. Your Family *

*Mark only one oval.*

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Very unhappy | | | | | | | | | | | | Very happy |

41. Your Friends *

*Mark only one oval.*

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Very unhappy | | | | | | | | | | | | Very happy |

42. Your Appearance (how you look) *

*Mark only one oval.*

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Very unhappy | | | | | | | | | | | | Very happy |

43. Your Life *

*Mark only one oval.*

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Very unhappy | | | | | | | | | | | | Very happy |

## You and your Feelings

Based on the Me and My Feelings Questionnaire ( Deighton, Tymms, Vostanis, Belsky, Fonagy, Brown, Martin, Patalay, & Wolpert, 2012)

44. 20. Remember, there are no right or wrong answers, just pick which is right for you. *

*Mark only one oval per row.*

| | Never | Sometimes | Always |
|---|---|---|---|
| I feel lonely | ◯ | ◯ | ◯ |
| I cry a lot | ◯ | ◯ | ◯ |
| I am unhappy | ◯ | ◯ | ◯ |
| I feel nobody likes me | ◯ | ◯ | ◯ |
| I worry a lot | ◯ | ◯ | ◯ |
| I have problems sleeping | ◯ | ◯ | ◯ |
| I wake up in the night | ◯ | ◯ | ◯ |
| I am shy | ◯ | ◯ | ◯ |
| I feel scared | ◯ | ◯ | ◯ |
| I worry when I am at school | ◯ | ◯ | ◯ |
| I get very angry | ◯ | ◯ | ◯ |
| I lose my temper | ◯ | ◯ | ◯ |
| I hit out when I am angry | ◯ | ◯ | ◯ |
| I do things to hurt people | ◯ | ◯ | ◯ |
| I am calm | ◯ | ◯ | ◯ |
| I break things on purpose | ◯ | ◯ | ◯ |

## Your Local Area

45. 21. On a scale of 0 to 10 (0 being not very safe and 10 being very safe), how safe do you feel playing in your area? *

*Mark only one oval.*

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Not very safe | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | Very safe |

46. 22a. From your house, can you walk to school?


*Mark only one oval.*

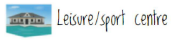
◯ Yes
◯ No

47. 22b. From your house, can you easily walk to a park?


*Mark only one oval.*

◯ Yes
◯ No

48. 22c. From your house, can you easily walk to a leisure centre/sports centre?




*Mark only one oval.*

◯ Yes
◯ No

49. 23. Are you happy with the area that you live in?

*Mark only one oval.*

◯ Yes
◯ No

24. If you could change something to make you and your friends healthier and happier, what would you change...

50. IN SCHOOL? *

51. OUT OF SCHOOL? *

Well done, you've completed the questionnaire.
Thank you!


Don't forget to press submit below!

**Online supplemental appendix 4: HAPPEN survey variable codebook**

| Exposures | HAPPEN Survey item | Responses | Analyses coding |
|---|---|---|---|
| **Ate breakfast** | *13. What did you eat for breakfast yesterday?* | *Nothing*<br>*Cereal*<br>*Snacks*<br>*Fruit*<br>*Toast*<br>*Cooked breakfast*<br>*Yoghurt* | Binary:<br>*1 = Cereal; Snacks; Fruit; Toast; Cooked breakfast; Yoghurt*<br>*0 = Nothing* |
| **Active travel to school** | *14. How did you get to school yesterday morning?* | *On the bus*<br>*In the car/taxi*<br>*Walked*<br>*On bike*<br>*Ran/jogged*<br>*Scooter*<br>*Skateboarded/rollerbladed* | Binary:<br>*1 = Walked; On bike; Ran/jogged; Scooter; Skateboarded/rollerbladed*<br>*0 = On the bus; In the car/taxi* |
| **Active travel from school** | *18. How did you get home yesterday?* | *On the bus*<br>*In the car/taxi*<br>*Walked*<br>*On bike*<br>*Ran/jogged*<br>*Scooter*<br>*Skateboarded/rollerbladed* | Binary:<br>*1 = Walked; On bike; Ran/jogged; Scooter; Skateboarded/rollerbladed*<br>*0 = On the bus; In the car/taxi* |
| **Toothbrush 2+ per day** | *20. How many times did you brush your teeth yesterday?* | *0 – 3* | Continuous:<br>*0 – 3* |
| **5+ fruit and veg** | *19. How many portions of fruit and vegetables did you eat yesterday?* | *0 – 8* | Continuous:<br>*0 – 8* |

**Online supplemental appendix 4: HAPPEN survey variable codebook**

| | | | |
|---|---|---|---|
| **Sleep 9+ hours** | *21. What time did you fall asleep last night* | *(30 min intervals)*<br>*7:00pm – 4:00am* | Continuous:<br>*Sleep hours calculated from item 21 and 22* |
| | *22. What time did you wake up this morning?* | *(30 min intervals)*<br>*5:00am – 9:00am* | |
| **Physically active 60+ mins every day previous 7 days** | *23. In the last 7 days, how many days did you do sports or exercise for at least 1 hour in total (This includes doing any activities or playing sports where your heart beat faster, you breathed faster and you felt warmer* | *0 days*<br>*1 – 2 days*<br>*3 – 4 days*<br>*5 – 6 days*<br>*7 days* | Ordinal:<br>*0 days*<br>*1 – 2 days*<br>*3 – 4 days*<br>*5 – 6 days*<br>*7 days* |
| **Sedentary/screen time 2 hours every day previous 7 days** | *24. In the last 7 days, how many days did you watch TV/play online games/use the internet etc. for 2 or more hours a day (in total)?* | *0 days*<br>*1 – 2 days*<br>*3 – 4 days*<br>*5 – 6 days*<br>*7 days* | Ordinal:<br>*0 days*<br>*1 – 2 days*<br>*3 – 4 days*<br>*5 – 6 days*<br>*7 days* |
| **Tired 7 days** | *25. In the last 7 days, how many days did you feel tired?* | *0 days*<br>*1 – 2 days*<br>*3 – 4 days*<br>*5 – 6 days*<br>*7 days* | Ordinal:<br>*0 days*<br>*1 – 2 days*<br>*3 – 4 days*<br>*5 – 6 days*<br>*7 days* |

**Online supplemental appendix 4: HAPPEN survey variable codebook**

| | | | |
|---|---|---|---|
| **Sugary snack 7 days** | *28. In the last 7 days, how many days did you eat at least one sugary snack (e.g. chocolate bar, sweets)* | *0 days*<br>*1 – 2 days*<br>*3 – 4 days*<br>*5 – 6 days*<br>*7 days* | Ordinal:<br>*0 days*<br>*1 – 2 days*<br>*3 – 4 days*<br>*5 – 6 days*<br>*7 days* |
| ***Participate in at least 3 out of school clubs*** | *31. How many times do you take part in a sports club OUTSIDE OF SCHOOL each week?* | *0 - 10* | Continuous:<br>*0 - 10* |
| ***Can ride a bike*** | *35. Can you ride a bike without stabilisers?* | *No*<br>*Yes* | Binary:<br>*1 = Yes*<br>*0 = No* |
| ***Can swim 25m*** | *36. Can you swim 25 metres without a float or armbands (This is 1 length of a standard swimming pool)* | *No*<br>*Yes* | Binary:<br>*1 = Yes*<br>*0 = No* |
| ***Age on 01/03/2020*** | *Decimal age on 1 March 2020* | Continuous | Continuous |
| ***Sex*** | *Sex* | *Girl*<br>*Boy* | Binary:<br>*0 = Girl*<br>*1 = Boy* |
| ***WIMD*** | *Welsh Index of Multiple Deprivation 2019* | | Coding framework from WIMD 2019[34] |

**Online supplemental appendix 5:** Full descriptive statistics table by tested for SARS-CoV-2 and tested positive for SARS-CoV-2.

| | | Tested for SARS-CoV-2 % (n) | Not tested for SARS-CoV-2 % (n) | Tested positive for SARS-CoV-2 % (n) | Tested negative/not tested (unknown) for SARS-CoV-2 % (n) |
|---|---|---|---|---|---|
| **LINKED DATA** | | | | | |
| **Sample** | | 39.1% (2,764) | 60.9% (4,298) | 8.1% (569) | 91.9% (6,493) |
| **Age at time of HAPPEN survey** | | 10.1 ± 0.8 | 9.9 ± 0.9 | 10.1 ± 0.8 | 9.9 ± 0.8 |
| **Age on 01/03/2020 (start of period of interest)** | | 10.6 ± 0.9 | 10.3 ± 1.1 | 10.6 ± 1.0 | 10.4 ± 1.0 |
| **Sex** | **Boy** | 49.3% (1,363) | 46.7% (2,005) | 44.3% (252) | 48.0% (3,116) |
| | **Girl** | 48.9% (1,352) | 51.8% (2,226) | 54.5% (310) | 50.3% (3,268) |
| | *Missing* | *1.8% (49)* | *1.5% (67)* | 1.2% (7) | 1.7% (109) |
| **WIMD 2019 quintiles** | **1 (most deprived)** | 24.3% (672) | 23.9% (1,025) | 28.5% (162) | 23.6% (1,535) |
| | **2** | 19.9% (551) | 19.02% (826) | 19.7% (112) | 19.5% (1,265) |
| | **3** | 16.5% (455) | 17.4% (748) | 17.6% (100) | 17.0% (1,103) |
| | **4** | 15.6% (431) | 15.8% (678) | 14.1% (80) | 15.9% (1,029) |
| | **5 (least deprived)** | 18.0% (497) | 16.8% (771) | 16.5% (94) | 17.3% (1,124) |
| | *Missing* | *5.7% (158)* | *7.0% (300)* | *3.7% (21)* | *6.7% (437)* |
| **HAPPEN SURVEY** | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| | | **Previous day** | | | |
| **Ate breakfast** | **Yes** | 93.0% (2,571) | 92.1% (3,797) | 93.4% (538) | 92% (6,012) |
| | **No** | 7% (193) | 7.3% (319) | 5.6% (31) | 7.3% (481) |
| | *Missing* | *0%* | *0%* | *0%* | *0%* |
| **Active travel to school** | **Yes** | 38.5% (1,065) | 39.8% (1,710) | 37.6% (214) | 39.4% (2,561) |
| | **No** | 61.5% (1,699) | 60.2% (2,588) | 62.4% (355) | 60.6% (3,932) |
| | *Missing* | *0%* | *0%* | *0%* | *0%* |
| **Active travel from school** | **Yes** | 43.0% (1,187) | 43.0% (1,846) | 42.4% (241) | 43.0% (2,792) |
| | **No** | 57.0% (1,577) | 57.0% (2,452) | 57.6% (328) | 57.0% (3,701) |
| | *Missing* | *0%* | *0%* | *0%* | *0%* |
| **Toothbrush continuous** | **0** | 3.3% (91) | 3.4% (146) | 1.9% (11) | 3.5% (227) |
| | **1** | 20.0% (552) | 21.0% (903) | 18.6% (106) | 20.6% (1,358) |
| | **2** | 67.1% (1,854) | 65.2% (2,802) | 69.6% (396) | 65.2% (4,294) |
| | **3** | 9.6% (265) | 10.3% (446) | 9.5% (54) | 10.0% (659) |
| | *Missing* | *0.1% (<5)* | *<0.1% (<5)* | *0.4% (<5)* | *<0.1%% (<5)* |
| **Fruit/veg portions (continuous)** | **0** | 14.3% (395) | 15.3% (657) | 12.5% (71) | 15.1% (981) |
| | **1** | 16.1% (445) | 17.4% (749) | 15.8% (90) | 17.0% (1,104) |
| | **2** | 17.7% (489) | 17.5% (754) | 19.5% (111) | 17.4% (1,132) |
| | **3** | 17.5% (484) | 16.5% (711) | 16.7% (95) | 16.9% (1,110) |
| | **4** | 12.7% (351) | 11.9% (510) | 13.5% (77) | 12.1% (784) |
| | **5** | 10.5% (291) | 10.6% (455) | 11.8% (67) | 10.4% (679) |
| | **6** | 4.5% (123) | 4.3% (186) | 2.8% (16) | 4.5% (293) |
| | **7** | 2.3% (63) | 2.1% (92) | 4.2% (24) | 2.0% (131) |
| | **8** | 4.5% (123) | 4.3% (184) | 3.2% (18) | 4.5% (289) |
| | *Missing* | 0% | 0% | *0%* | *0%* |

| | | | | | |
|---|---|---|---|---|---|
| **Sleep hours** | | 9.4 ± 1.6 | 9.4 ± 1.6 | 9.4 ± 1.6 | 9.4 ± 1.6 |
| **Number of days physically active ≥ 60 minutes** | **0** | 6.5% (179) | 7.9% (339) | 4.0% (23) | 7.6% (495) |
| | **1-2 days** | 27.9% (772) | 29.0% (1,246) | 27.8% (158) | 28.7% (1,860) |
| | **3-4 days** | 27.5% (761) | 26.2% (1,128) | 30.9% (176) | 26.4% (1,712) |
| | **5-6 days** | 18.3% (505) | 17.0% (731) | 18.1% (103) | 17.5% (1,133) |
| | **7 days** | 19.8% (557) | 19.9% (854) | 19.2% (109) | 19.9% (1,292) |
| | *Missing* | *0%* | *0%* | *0%* | *0%* |
| **Number of days sedentary/screen time ≥ two hours** | **0** | 5.2% (144) | 6.1% (262) | 5.5% (31) | 5.8% (375) |
| | **1-2 days** | 24.2% (674) | 23.5% (1,011) | 24.8% (141) | 23.8% (1,544) |
| | **3-4 days** | 21.7% (599) | 20.6% (886) | 21.1% (120) | 21.0% (1,365) |
| | **5-6 days** | 14.0% (386) | 13.8% (593) | 13.9% (79) | 13.9% (900) |
| | **7 days** | 34.8% (961) | 36.0% (1,546) | 34.8% (198) | 35.6% (2,309) |
| | *Missing* | *0%* | *0%* | *0%* | *0%* |
| **Number of days tired** | **0** | 21.0% (582) | 21.0% (903) | 19.2% (109) | 21.2% (1,376) |
| | **1-2 days** | 32.4% (895) | 32.0% (1,377) | 35.7% (203) | 31.9% (2,069) |
| | **3-4 days** | 17.6% (487) | 17.5% (754) | 18.8% (107) | 17.5% (1,134) |
| | **5-6 days** | 10.0% (276) | 9.3% (399) | 10.5% (60) | 9.5% (615) |
| | **7 days** | 19.0% (524) | 20.1% (865) | 15.8% (90) | 20.0% (1,299) |
| | *Missing* | *0%* | *0%* | *0%* | *0%* |
| **Number of days sugary snack** | **0** | 6.5% (179) | 7.7% (332) | 6.3% (36) | 7.3% (475) |
| | **1-2 days** | 34.9% (964) | 32.7% (1,407) | 35.0% (199) | 33.5% (2,172) |
| | **3-4 days** | 25.3% (698) | 26.7% (1,146) | 25.1% (143) | 26.2% (1,701) |
| | **5-6 days** | 13.4% (371) | 12.0% (515) | 15.3% (87) | 12.3% (799) |
| | **7 days** | 20.0% (552) | 20.9% (898) | 18.3% (104) | 20.7% (1,346) |

| | | | | | |
|---|---|---|---|---|---|
| | *Missing* | *0%* | *0%* | *0%* | *0%* |
| | | **General** | | | |
| **Number of out of school clubs** | **0** | 27.7% (766) | 32.3% (1,387) | 25.1% (143) | 31.0% (2,010) |
| | **1** | 17.9% (495) | 16.9% (726) | 16.0% (91) | 17.4% (1,130) |
| | **2** | 16.0% (443) | 15.1% (650) | 14.9% (85) | 15.5% (1,008) |
| | **3** | 11.1% (308) | 10.4% (446) | 13.3% (76) | 10.4% (678) |
| | **4** | 7.4% (204) | 7.3% (313) | 7.6% (43) | 7.3% (474) |
| | **5** | 6.2% (171) | 5.8% (251) | 5.8% (33) | 6.0% (389) |
| | **6** | 3.4% (95) | 2.5% (109) | 5.1% (29) | 2.7% (175) |
| | **7** | 3.3% (91) | 2.5% (107) | 5.1% (29) | 2.6% (169) |
| | **8** | 1.1% (29) | 0.8% (33) | 1.8% (10) | 0.8% (52) |
| | **9** | 0.9% (24) | 0.7% (32) | 1.2% (7) | 0.8% (49) |
| | **10** | 3.9% (107) | 4.0% (174) | 3.3% (19) | 4.0% (262) |
| | *Missing* | *1.1% (31)* | *1.6% (70)* | *0.7% (<5)* | *1.5% (97)* |
| **Can ride a bike** | **Yes** | 88.8% (2,444) | 86.0% (3,696) | 91.4% (520) | 86.7% (5,641) |
| | **No** | 11.2% (309) | 14.0% (602) | 8.6% (49) | 13.3% (862) |
| | *Missing* | *0%* | *0%* | *0%* | *0%* |
| **Can swim 25m** | **Yes** | 78.9% (2,180) | 72.9% (3,134) | 80.3% (457) | 74.8% (4,857) |
| | **No** | 21.1% (584) | 27.1% (1,164) | 19.7% (112) | 25.2% (1,636) |
| | *Missing* | *0%* | *0%* | *0%* | *0%* |

**Online supplemental appendix 6:**
Multivariable logistic regression model of health behaviour markers and probability of PCR-test without confounders.

| PCR tested for SARS-CoV-2 (n=6,958, $R^2$=0.01) | OR | *p* value | 95% CI |
|---|---|---|---|
| **Ate breakfast** | 1.05 | 0.632 | 0.87 – 1.27 |
| *Reference: did not eat breakfast* | 1.00 | | |
| **Active travel to school** | 0.92 | 0.238 | 0.80 – 1.06 |
| *Reference: did not active travel to school* | 1.00 | | |
| **Active travel from school** | 1.08 | 0.273 | 0.94 – 1.24 |
| *Reference: did not active travel from school* | 1.00 | | |
| **Number of fruit/vegetable portions** | 1.00 | 0.941 | 0.98 – 1.03 |
| **Number of times teeth brushed** | 0.97 | 0.474 | 0.90 – 1.05 |
| **Sleep hours** | 0.99 | 0.654 | 0.96 – 1.02 |
| *Reference: 0 days physically active ≥60 mins (previous seven days)* | 1.00 | | |
| **1-2 days physically active ≥ 60 mins** | 1.12 | 0.276 | 0.91 – 1.38 |
| **3-4 days physically active ≥ 60 mins** | 1.14 | 0.221 | 0.92 – 1.42 |
| **5-6 days physically active ≥ 60 mins** | 1.17 | 0.177 | 0.93 – 1.47 |
| **7 days physically active ≥ 60 mins** | 1.09 | 0.475 | 0.87 – 1.37 |
| *Reference: 0 days sedentary ≥two hours (previous seven days)* | 1.00 | | |
| **1-2 days sedentary ≥ two hours** | 1.16 | 0.209 | 0.92 – 1.46 |
| **3-4 days sedentary ≥ two hours** | 1.18 | 0.166 | 0.93 – 1.49 |
| **5-6 days sedentary ≥ two hours** | 1.15 | 0.275 | 0.90 – 1.47 |
| **7 days sedentary ≥ two hours** | 1.14 | 0.256 | 0.91 – 1.44 |
| *Reference: 0 days felt tired (previous seven days)* | 1.00 | | |
| **1-2 days felt tired** | 0.98 | 0.791 | 0.86 – 1.13 |
| **3-4 days felt tired** | 0.99 | 0.881 | 0.84 – 1.16 |
| **5-6 days felt tired** | 1.04 | 0.667 | 0.86 – 1.26 |
| **7 days felt tired** | 0.97 | 0.730 | 0.83 – 1.14 |
| *Reference: 0 days consumed sugary snack (previous seven days)* | | | |
| **1-2 days consumed sugary snack** | 1.21* | 0.062 | 0.99 – 1.49 |
| **3-4 days consumed sugary snack** | 1.08 | 0.489 | 0.87 – 1.33 |
| **5-6 days consumed sugary snack** | 1.29** | 0.034 | 1.02 – 1.63 |
| **7 days consumed sugary snack** | 1.12 | 0.314 | 0.90 – 1.39 |
| **Number of out of school clubs participation** | 1.02 | 0.121 | 1.00 to 1.04 |
| **Can ride a bike** | 1.16* | 0.064 | 0.99 – 1.35 |
| *Reference: cannot ride a bike* | 1.00 | | |
| **Can swim 25m** | 1.30** | < 0.001 | 1.15 – 1.46 |

| | |
|---|---|
| *Reference: cannot swim 25m* | 1.00 |

OR: Odds Ratio; 95% CI: 95% confidence intervals; p<0.05**, p<0.1*. See online supplemental appendix 4 for variable codebook.

Multivariable logistic regression model of health behaviour markers and probability of PCR-test positive without confounders.

| **PCR test positive for SARS-CoV-2 (n=6,958, $R^2$=0.01)** | **OR** | ***p* value** | **95% CI** |
|---|---|---|---|
| **Ate breakfast** | 1.30 | 0.170 | 0.89 – 1.91 |
| *Reference: did not eat breakfast* | 1.00 | | |
| **Active travel to school** | 0.91 | 0.451 | 0.71 – 1.17 |
| *Reference: did not active travel to school* | 1.00 | | |
| **Active travel from school** | 1.07 | 0.614 | 0.83 – 1.36 |
| *Reference: did not active travel from school* | 1.00 | | |
| **Number of fruit/vegetable portions** | 0.99 | 0.574 | 0.94 – 1.03 |
| *Reference: 0 fruit/vegetable portions* | 1.00 | | |
| **Number of times teeth brushed** | 1.07 | 0.385 | 0.92 – 1.24 |
| *Reference: did not brush teeth* | 1.00 | | |
| **Sleep hours** | 0.97 | 0.266 | 0.92 – 1.02 |
| *Reference: 0 days physically active ≥60 mins (previous seven days)* | 1.00 | | |
| **1-2 days physically active ≥ 60 mins** | 1.71 | 0.023 | 1.08 – 2.73 |
| **3-4 days physically active ≥ 60 mins** | 1.87 | 0.009 | 1.17 – 2.99 |
| **5-6 days physically active ≥ 60 mins** | 1.61 | 0.059 | 0.98 – 2.63 |
| **7 days physically active ≥ 60 mins** | 1.49 | 0.117 | 0.91 – 2.43 |
| *Reference: 0 days sedentary ≥two hours (previous seven days)* | 1.00 | | |
| **1-2 days sedentary ≥ two hours** | 1.03 | 0.877 | 0.68 – 1.57 |
| **3-4 days sedentary ≥ two hours** | 1.00 | 0.983 | 0.66 – 1.54 |
| **5-6 days sedentary ≥ two hours** | 1.01 | 0.958 | 0.65 – 1.59 |
| **7 days sedentary ≥ two hours** | 1.10 | 0.660 | 0.72 – 1.66 |
| *Reference: 0 days felt tired (previous seven days)* | 1.00 | | |
| **1-2 days felt tired** | 1.21 | 0.125 | 0.95 – 1.55 |
| **3-4 days felt tired** | 1.17 | 0.278 | 0.88 – 1.55 |
| **5-6 days felt tired** | 1.21 | 0.273 | 0.86 – 1.69 |
| **7 days felt tired** | 0.92 | 0.600 | 0.69 – 1.24 |
| *Reference: 0 days consumed sugary snack (previous seven days)* | 1.00 | | |
| **1-2 days consumed sugary snack** | 1.14 | 0.499 | 0.78 – 1.67 |
| **3-4 days consumed sugary snack** | 1.03 | 0.873 | 0.70 – 1.53 |
| **5-6 days consumed sugary snack** | 1.38 | 0.131 | 0.91 – 2.11 |
| **7 days consumed sugary snack** | 1.04 | 0.867 | 0.69 – 1.56 |
| **Number of out of school clubs participation** | 1.05 | 0.007 | 1.01 – 1.09 |
| **Can ride a bike** | 1.40 | 0.032 | 1.03 – 1.92 |

| | | | |
|---|---|---|---|
| *Reference: cannot ride a bike* | 1.00 | | |
| **Can swim 25m** | 1.16 | 0.207 | 0.92 – 1.45 |
| *Reference: cannot swim 25m* | 1.00 | | |