EVENTS RESEARCH PROGRAMME: NATIONAL STATISTICAL ANALYSIS PLAN (SAP)

1. Administrative information

1.1. Title, registration, versions and revisions

Full study title	Descriptive and feasibility study of risk-mitigation with SARS-CoV-2 antigen rapid lateral flow testing to reopen live events
Short study title	Mitigating COVID-19 transmission risk at events with rapid testing
Study protocol version	PHE 1.5 (23 April 2021)
SAP version	2.6 (17 May 2021)

1.2. Roles and responsibilities

Contributors and roles	Anna Trelfa[1]: Lead author/Analyst (public health) Jenifer Smith[1]: Co-Principal Investigator (public health) Alex Cockburn[1]: Analyst (public health) Chris Cheyne[2]: Analyst (statistics) Girvan Burnside[2]: Analyst (statistics) Iain Buchan[2]: Co-Principal Investigator (public health/statistics) Marta Garcia-Fiñana[2]: Reviewer/contributor (statistics) Simon Maskell[2]: Reviewer/contributor (modelling) John Edmunds[3]: Reviewer/contributor (modelling)
Affiliations	1. Public Health England 2. University of Liverpool 3. London School of Hygiene and Tropical Medicine

2. Introduction

2.1. Background and rationale

The Events Research Programme (ERP) is exploring the use of COVID-19 risk-mitigation strategies to support reopening a variety of events and venues, as part of the Government's Roadmap for COVID-19 recovery. The Roadmap describes plans to explore when and how events with large crowd sizes, less social distancing, or in settings where transmission risk is high without mitigation, may be able to return safely.

Sports, entertainment, social and business functions that are currently restricted carry unknown risks of SARS-CoV-2 transmission, particularly regarding the latest variants. So, there is a need to mitigate the likely increases in transmission due to social mixing at such events/venues. Currently a range of non-pharmaceutical interventions (NPIs) are used to mitigate risk, but these impact on the participant experience and satisfaction as well as the commercial viability of events whilst the relative contribution of such interventions to controlling the transmission of COVID-19 is not known.

The ERP will use events incorporating research studies in April and May 2021 to build evidence of the risks associated with coronavirus transmission, the characteristics of events and surrounding activities, and the most effective steps for reducing these risks. The evidence from these studies will be used to inform the Government's decisions around Step 4 of the Roadmap—removal of all legal limits on social contact, reopening of any remaining premises, including nightclubs, and easing of

restrictions on large events, performances, weddings, and other life events—and will help shape Government policy to bring about the phased return of fuller audiences to venues and events. Step 4 of the Roadmap will take place no earlier than 21 June.

The ERP will draw findings from the evaluation of a series of scientifically instrumented events run by stakeholders across a range of venue types (including indoor and outdoor) and sectors (arts, creative industries, business, and sport events). Building the evidence base across a range of settings will enable more tailored mitigation measures to be considered rather than one-size fits all approaches (such as current capacity caps). The results of the testing evaluation will be aligned with the outputs from the concurrent research studies on ventilation, crowd behaviour and social distancing to inform risk mitigation measures for Stage 4 of the Roadmap.

This Statistical Analysis Plan (SAP) focuses on the analyses required to answer a set of national research questions, as detailed in section 2.3. Descriptive analyses for each of the individual events will also be reported here. Additional research questions and analysis for events held in Liverpool are detailed elsewhere (see Re-Liv-E protocol and SAP).

2.2. Science Board Statement

The following text is extracted from the Science Board Statement (March 2021), which summarises the evidence that the ERP aims to generate.

"1. At current levels of prevalence, very large numbers of attendees and comparable comparison groups would be required in order to estimate epidemiologically meaningful differences in risk of transmission of SARS-CoV-2 across event types and mitigation measures (Science Framework 16.03.21).

2. The initial events within the Events Research Programme (ERP) are currently insufficient in scale, scope and study designs to generate any direct evidence based on transmission data on how opening events might be done to mitigate risks of transmission.

3. The initial events within ERP will, however, provide evidence on the potential effectiveness of i. ventilation systems, assessed using CO2 measurements as a proxy for flow rate per person; and ii. organisation of events and venue design assessed using video footage of attendee behaviour (see 5c below).

4. By pooling data across events, the ERP might provide an indication of risk of transmission by comparing rates of infection in attendees post-event, with rates of infection based on surveillance data in similar populations. Any such estimate will be extremely tentative given (a) none of the planned studies includes a comparison group and (b) post-event test results may only be available on a small and unrepresentative sample of attendees.

5. The ERP will generate important evidence to inform a further phase of the ERP in which direct evidence of mitigating transmission could be generated including:

a. feasibility and acceptability of pre- and post-event testing, including:
- *proportion (and demographic representativeness) of those expressing interest in attending an event who undergo testing and attend;*
- *proportion of those attending who return a post-event test kit (and demographic representativeness)*

b. impact of incentives on post-event test return rates

c. feasibility of using video footage to measure attendees' behaviour at events:

- *movement (to estimate aerobic levels)*
- *interpersonal distances and interactions*
- *wearing of face coverings*

6. The evidence generated by ERP will be supplemented by results from an ongoing rapid review of evidence, being undertaken by PHE, and by summaries of mass gatherings, prepared by the International Comparators Joint Unit, FCDO."

2.3. Aim and objectives

2.3.1. Aim, objectives and research questions

The aim of this study is to provide evidence on the feasibility of pre-event rapid antigen testing with lateral flow devices (LFD) for SARS-CoV-2 in mitigating the risk of COVID-19 transmission amongst spectators, participants or audiences at cultural and sporting events, instrumented with and pre- and post-event PCR testing.

The objectives of this study are to:

1. Estimate the PCR test uptake rate per event and pooled across all events, and identify individual- and event-level factors that are associated with PCR test uptake rate.
2. Estimate the rate of index cases and potential secondary cases in eventgoers per event and pooled across all events, and describe the variation by individual- and event-level factors.
3. Identify event-related clusters using genomic sequencing data.
4. Understand the data flows required to facilitate control of transmission of SARS-CoV-2 arising from attendance at events.
5. Undertake a gap analysis of events-related contact tracing, to identify what level of detail is currently collected by the national Contact Tracing and Advisory Service (CTAS) and the opportunities for collecting specific events-related information from cases.

The study research questions are:

1. What proportion of those who attend the events:
    a. complete their LFD test at an asymptomatic test site (ATS) as per protocol (for events dated 17 April-11 May),
    b. return a completed pre-event home test PCR kit,
    c. return a completed post-event home test PCR kit, and
    d. return both completed pre-event and post-event home test PCR kits?

Descriptive tables will be used to show PCR uptake proportions stratified by sociodemographic factors (e.g. age, sex, ethnicity, deprivation), vaccination status, previous infection status, indoor/outdoor event, seated/standing/mixed event, mean crowd density and method of accessing PCR testing. Generalised linear modelling will be applied per event to identify individual characteristics/factors that are associated with pre- and post-event PCR uptake. Pooled data analysis will be based on a generalised linear mixed effects model that accounts for the hierarchical nature of the data.

2. What proportion of those who attend the events:
    a. receive a PCR positive test result indicative of them attending the event already infected (i.e. index cases), and
    b. receive a post-event PCR positive test result and is suspected to be a secondary case (e.g. pre-event PCR negative test result)?

Descriptive tables will summarise proportions of positive cases stratified by sociodemographic factors (e.g. age, sex, ethnicity, deprivation), vaccination status, previous infection status,

indoor/outdoor event, seated/standing/mixed event and mean crowd density. To distinguish between a pre-event PCR and a post-event PCR, we will use the following definitions:

Pre-event: Specimen date 2 days before to 3 days following the event:	A **positive** result = **index** case
Post-event: Specimen date 4-7 days following the event:	A **positive** result = **putative secondary**
Specimen 8+ days after the event	A **positive** result = **not directly related**

We will use cycle threshold (Ct) values of PCR results to estimate dates of infection and explore secondary case status further, taking other contextual information into account. A sensitivity analysis will account for missing data on PCR positivity using multiple imputation. For cases in which only a single PCR result is available the Ct value and specimen date will be used to assign index, putative secondary, or not-directly-related case status probabilistically, using a mathematical model of Ct-value progression with time since infection (Quilty et al. Lancet Public Health, 2021). Descriptions of potential infector-infectee pairs will be provided, using further contextual information such as whether index and putative secondary cases lived at the same address, or shared transport to the venue. This additional information will not be used to assign case status in the base case.

Using whole genome sequencing of positive home test PCR kits:

c. how many likely secondary (event-transmitted) cases were detected with post-event PCR tests and are linked by their genomic sequencing results, and
d. can genomic sequencing of pre-event PCR positive tests identify a linked index case?

If enough high-quality data is available (i.e. assuming sufficient matching between ticketing and testing data, as well as an adequate PCR return rate), the following research questions will also be explored:

i. the difference in post-event incidence rate of SARS-CoV-2 infection between events with no index cases identified and events where index cases were identified,
ii. associations between post-event PCR positivity, and individual- and event-level characteristics, in those who had a negative pre-event PCR test result, and
iii. associations between genomically-confirmed cluster size and event-level characteristics.

2.3.2. Scope

This SAP will be the guiding document for the analyses that will be conducted in the national ERP LFD feasibility study for objectives 1-3 and should be considered alongside the Re-Liv-E SAP.

3. Study methods

3.1. General study design and plan

There are 4 indoor events at small to medium sized venues and 6 outdoor events, as detailed in the table that follows. The venues will host a range of sporting and cultural events with audiences of varying size and capacity for the venues. There are 3 multi-day events.

Table 1. Summary of ERP phase 1 events.

Date	Event	Location	Attendees	Event type
18 April	FA Cup Semi-final	Wembley, London	4,000	Outdoor seated
17 April -3 May	Snooker World Championships	Crucible Theatre, Sheffield	<1,000/day	Indoor seated

Date	Event	Location	Attendees	Event type
25 April	Carabao Cup	Wembley, London	8,000	Outdoor seated
28 April	Good Business Event	Arena, Liverpool	300	Indoor mixed – open/seated
30 April -1 May	Circus Nightclub	Bramley Moore Dock, Liverpool	3,000 and 4,000	Indoor open
2 May	Live National Outdoor Music Event	Sefton Park, Liverpool	6,000	Outdoor unstructured
11 May	Brit Awards	O2, London	5,000	Indoor seated (mixed styles)
15 May	FA Cup Final	Wembley, London	21,000	Outdoor seated
15 May	Mass Participation Run	Kempton Park, Surrey	18,000	Outdoor unstructured
		Maximum sample size:	c. 86,300	

This study aims to compare SARS-CoV-2 PCR test positivity pre- and post-attendance at an ERP event.

Every eventgoer will undertake an LFD test in the 36 hours prior to the event. Only those able to demonstrate a negative test result, and declaring no symptoms, will be permitted entry to the event. For all events, other than the FA Cup Final and the Mass Participation Run, eventgoers will be required to undertake their pre-event LFD test at an ATS. For the FA Cup Final, 9,000 eventgoers attending the FA Cup Final will be randomised to either home LFD test or ATS LFD test, with the remaining 12,000 spectators required to undertake their pre-event LFD test at an ATS. For the Mass Participation Run, eventgoers will be given the choice of whether to undertake their pre-event LFD test at an ATS or at home.

Eventgoers at the Snooker World Championships attending multi-day events on successive days, will be asked to take an LFD test at an ATS every 3 days, and can only participate in the study if all tests are negative. If an LFD test is positive, the eventgoer must not attend the event and, in accordance with national guidance, they must self-isolate and take a confirmatory PCR test. To assess the transmission of COVID-19 at the event itself and the effectiveness of pre-event LFD testing in identifying those infected with SARS-CoV-2, eventgoers will also be asked to provide a home PCR test on the day of the event and a home PCR test five days after the event.

All positive PCR tests with a Ct value of less than 30 will be genome sequenced to identify clusters of SARS-CoV-2 infections, indicating potential transmission at the event. Note that due to lab process issues, a maximum of 89% of cases detected from the events that pre-date 30 April will undergo genomic sequencing; this will be closer to 100% for events occurring on or after 30 April (individual test Ct values permitting).

The event organisers will send a list of all eventgoers who attended the event to PHE, using secure transfer methods. Eventgoer information (self-reported name, date of birth, sex and full address) will be linked to NHS number using the Demographic Batch Service. NHS number will be used to link to the Pillar 2 testing dataset, for the time-period from 36 hours prior to 7 days following the event. A 7-day post-event cut-off will be used to capture those who might receive a positive result earlier (e.g. if they become symptomatic) and those who might return their sample late. Genomic sequencing data from the final events will take longer to come through due to lab processing times – expected processing time upper limit of 9 days – therefore, a further data look-up of PCR-confirmed cases in eventgoers against sequencing data will be necessary at a later timepoint.

If an NHS number cannot be matched via the Demographic Batch Service, probabilistic data matching techniques will be used to attempt to link eventgoer information to the Pillar 2 testing dataset. Eventgoer data linkage will be performed with other datasets to allow the use of socio-demographic and other data (such as ethnicity, vaccination and deprivation data).

Any positive tests will be reported through Test and Trace and contact tracing undertaken to ascertain detail of activity during the day of the event including travel, seating and activity at the venue. Local contact tracing hubs may pick up some of this work, including the Cheshire and Merseyside hub for Liverpool events.

Data will be obtained, processed, stored and analysed in accordance with information governance protocols. Personal identifiers will not be stored, and data will be analysed using unique identifiers attached to record level data. Results will be published in anonymised format.

3.2. Sample size

The study population is around 86,300 people attending one or more of the events comprising the ERP phase 1. The analyses are mainly descriptive in nature, and no formal power size calculation is presented. The main focus is on the feasibility of testing (matching between ticketing and testing data and PCR return rate) and its utility for outbreak management and forward events planning under different epidemic state scenarios.

3.3. Timing of final analysis

The final analysis will be conducted after receipt of all participant and genomic sequencing information.

This statistical analysis plan will be added to the Open Science Framework, before closure of the database and before analyses are conducted.

3.4. Timing of outcome assessments

Eventgoers have been asked to return a completed post-event home test PCR kit on day 5 after the event, however PCR test data from days 1-7 will be used to assess the outcomes, to capture any early or late returns.

4. Study population

4.1. Eligibility

Event organisers will invite interest in the event from the general public. Note that for some events (e.g. the Brits, football matches), event organisers will only invite from a select pool (e.g. key workers, or football fans). Eventgoers will purchase a ticket, and only those who consent to participation in the research and who can provide evidence of a negative LFD result will be permitted entry to the event. For the events in Liverpool, eventgoers will also have to complete a pre-event questionnaire before ticket can be purchased and will be asked to complete a post-event questionnaire. Eventgoers at multi-day Snooker sessions (attending events over more than three consecutive days) will be required to provide a negative LFD result every three days to continue their attendance.

4.1.1. Inclusion criteria

- Consent to participation in the research
- Age ≥16 years
- Negative LFD result in the 36 hours preceding the event start-time/every three days if attending a multi-day event
- Free from symptoms of COVID-19 (declared at venue admittance)

- Not isolating as a COVID-19 case, or a contact of a case

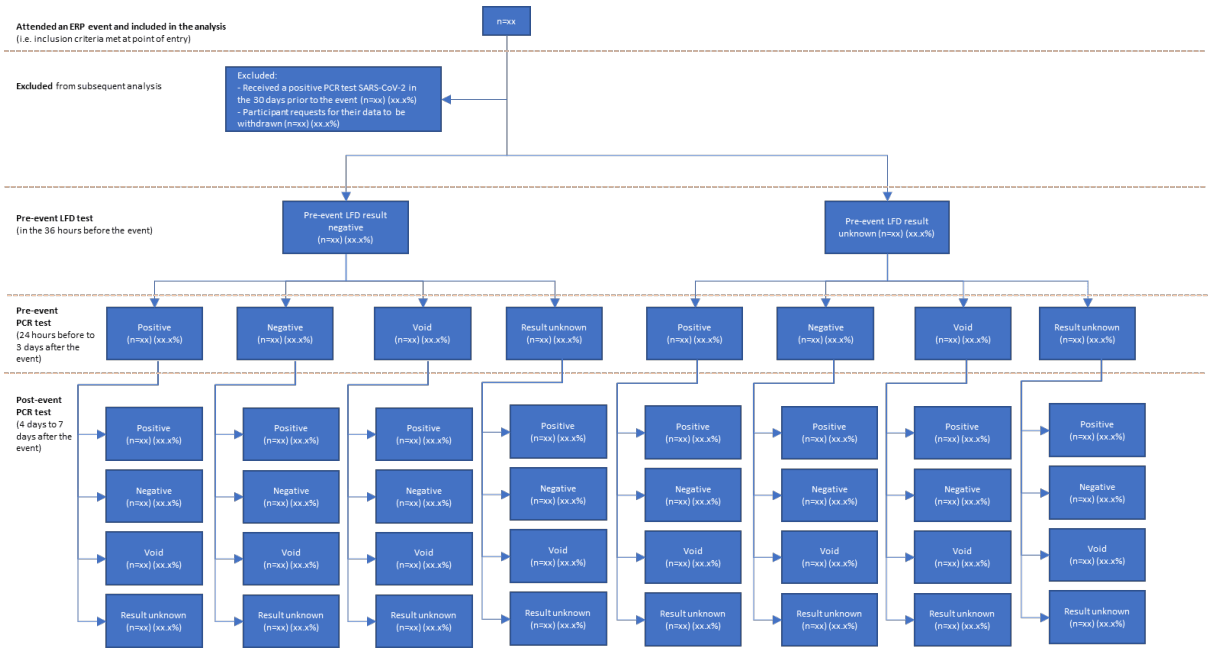
4.1.2. Exclusion criteria

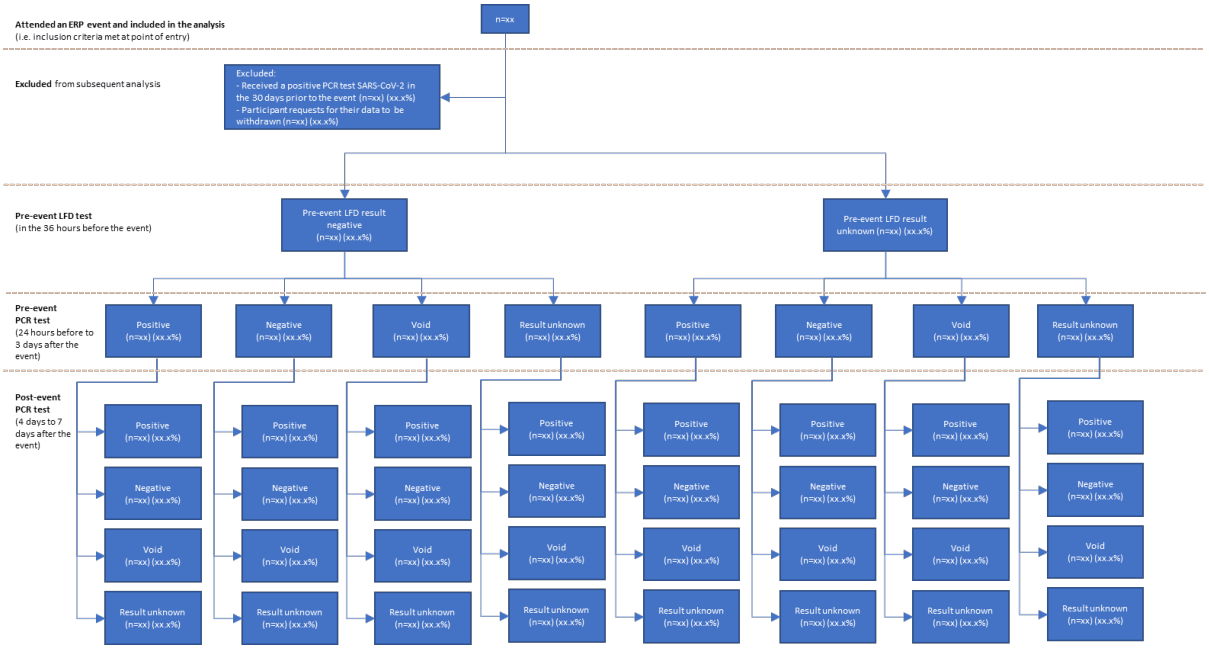
- Age < 16 years
- Has not consented to participation in the research
- Unable to provide proof of a pre-event negative LFD result at entry to the event
- Experiencing symptoms of COVID-19
- Within their self-isolation period, either as a COVID-19 case or a contact of a case
- Received a positive PCR test SARS-CoV-2 in the 30 days prior to the event – note that this criterion is applied after attendance at the event to avoid sending confusing messages to eventgoers. Given the very low current prevalence, this is likely to impact only a small number of the sample.

4.2. Recruitment and withdrawal

The flow of study participants is illustrated below. The first box contains all eventgoers who we know attended the event. 'Pre-event LFD result negative' includes all eventgoers for which a negative LFD test result in the 36 hours pre-event could be matched via NHS number or via fuzzy matching on demographic details. 'Pre-event LFD result unknown' includes all eventgoers for which a negative LFD test result in the 36 hours pre-event could not be matched via NHS number or probabilistic methods. Note that for the purpose of the analysis, it is assumed that all eventgoers did receive a pre-event negative LFD test result in the 36 hours prior to the event, as proof of this was required as a condition of entry. 'Result unknown' for both pre-event and post-event PCR tests includes all eventgoers for which a PCR test result could not be matched via NHS number or demographic details.

Figure 1. Flow diagram for pooled ERP participants, continued over the page.





4.3. Baseline participant characteristics

4.3.1. Collected baseline participant characteristics

The study has been designed to collect a set of socio-demographic and clinical variables per participant. Table 2 provides an overview of all participant variables.

Table 2. Participant variables captured in the national ERP dataset.

Variable	Source	Baseline	Every 3 days*	Post-event (days 1-7)
Participant ID	Study team	X		
Age	Derived	X		
Sex	SGSS	X		
Postcode	SGSS	X		
Ethnicity	SGSS	X		
IMDDecile	PHE COVID-19 Datastore	X		
VaccinationStatus	NHSX	X		
PrevCase2021	SGSS	X		
RegionalPrevRate	PHE Power BI COVID-19 Situational Awareness Explorer	X		
LFDSite	SGSS	X	X	
LFDDate	SGSS	X	X	
LFDTime	SGSS	X	X	
LFD36Hours	Derived	X	X	
LFDResult	SGSS	X	X	
PCRDate_D0	SGSS	X		
PCRResult_D0	SGSS	X		
PCRCtValue_D0	SGSS	X		
PCRSeq_D0	SGSS	X		
PCRDate_FU	SGSS			X
PCRResult_FU	SGSS			X
PCRCtValue_FU	SGSS			X
PCRSeq_FU	SGSS			X
PostEventPCRDay	Derived			X
PCRDelivery	DHSC	X		

*for eventgoers at multi-day Snooker sessions who attend events over more than three consecutive days.

4.3.2. Descriptive summary of baseline participant characteristics

We will describe general participant characteristics in a baseline characteristics summary table, as demonstrated in Table 3. This will include pooled data and characteristics split by the contributing events. Note that events have been selected to ensure greatest population coverage, therefore between-group variation is expected.

Table 3. Descriptive summary of participants who attended ERP events.

	FA Cup Semi-final	Snooker World Championships	Carabao Cup	Good Business Event	Circus Nightclub	Live National Outdoor Music Event	Brit Awards	FA Cup Final	Mass Participation Run	Pooled
No. of participants	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx
Age										
n	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx
n missing	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx
mean (SD)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)
median (Q1, Q3)	xx.x (xx.x-xx.x)	xx.x (xx.x-xx.x)	xx.x (xx.x-xx.x)	xx.x (xx.x-xx.x)	xx.x (xx.x-xx.x)	xx.x (xx.x-xx.x)	xx.x (xx.x-xx.x)	xx.x (xx.x-xx.x)	xx.x (xx.x-xx.x)	xx.x (xx.x-xx.x)
range (min, max)	xx.x-xx.x	xx.x-xx.x	xx.x-xx.x	xx.x-xx.x	xx.x-xx.x	xx.x-xx.x	xx.x-xx.x	xx.x-xx.x	xx.x-xx.x	xx.x-xx.x
Sex										
n	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx
n missing	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx
male, n (%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
female, n (%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)

	FA Cup Semi-final	**Snooker World Championships**	**Carabao Cup**	**God Business Event**	**Circus Nightclub**	**Live National Outdoor Music Event**	**Brit Awards**	**FA Cup Final**	**Mass Participation Run**	**Pooled**
Ethnicity										
n	xx	xx	xx	xx	xx	xx	xx	xx	xx	xx
n missing	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
Black, African, Black British, or Caribbean	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
Asian or Asian British	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
Another ethnic group	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
Mixed or multiple ethnic group	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
White	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
Prefer not to say	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
Index of Multiple Deprivation decile										
1	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
2	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
3	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
4	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
5	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
6	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
7	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
8	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
9	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
10	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)

	FA Cup Semi-final	Snooker World Championships	Carabao Cup	Good Business Event	Circus Nightclub	Live National Outdoor Music Event	Brit Awards	FA Cup Final	Mass Participation Run	Pooled
	xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%)	xx (xx.x%) xx (xx.x%) xx (xx.x%)	xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%)	xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%)	xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%)	xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%)	xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%)	xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%)	xx (xx.x%) xx (xx.x%) xx (xx.x%)	xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%) xx (xx.x%)
COVID-19 test positive in 2021 n n missing yes, n (%) no, n (%)	xx xx xx (xx.x%) xx (xx.x%)	xx xx xx (xx.x%) xx (xx.x%)	xx xx xx (xx.x%) xx (xx.x%)	xx xx xx (xx.x%) xx (xx.x%)	xx xx xx (xx.x%) xx (xx.x%)	xx xx xx (xx.x%) xx (xx.x%)	xx xx xx (xx.x%) xx (xx.x%)	xx xx xx (xx.x%) xx (xx.x%)	xx xx xx (xx.x%) xx (xx.x%)	xx xx xx (xx.x%) xx (xx.x%)
Vaccinated n n missing 1st dose, n (%) 2nd dose, n (%) none, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)

4.4. Event-level variables

4.4.1. Collected event characteristics

The following event characteristics will be collected for events:

- Event type (provided by AIRBODS study team):
    - indoor – poorly ventilated;
    - indoor – mechanically ventilated to COVID guidance standards;
    - indoor – naturally ventilated/low ventilation;
    - indoor – naturally ventilated/high ventilation;
    - outdoors – sheltered;
    - outdoors
- Audience (provided by DCMS):
    - seated;
    - mixed;
    - standing
- Mean crowd density (persons/m^2) (provided by Movement Strategies study team)
- Method of accessing PCR testing (provided by DHSC):
    - self-request;
    - pick-up from pre-event LFD ATS;
    - automatically posted

4.5. Assumed confounding factors

4.5.1. Measured confounders

PCR positivity is assumed to be confounded by the following variables, which we are collecting data on to enable stratified analyses:

- Age: protective – many of the older eventgoers will have been vaccinated.
- Vaccination status: protective.
- Ethnicity: Black, Asian and Minority Ethnic communities experience more infections (contributing either to an increased risk of infection or a level of protection due to population immunity) and more severe consequences.
- Deprivation: residents of more deprived communities experience more infections (contributing either to an increased risk of infection or a level of protection due to population immunity) and more severe consequences.
- Recent infection: recent positive test may protect against new infection.

4.5.2. Unmeasured/residual confounders

Individual-level data on the following potential confounding factors are not being collected for eventgoers, aside from those marked with an asterisk which are being collected for Liverpool events only:

- Pre- or post-event activities*
- Mode of travel to and from the event*
- Occupation
- Household size
- Recent exposure as a contact of a confirmed case

5. Analysis

5.1. Outcome definitions and analysis methods

1a. What proportion of those who attend the event complete their lateral flow device test at an asymptomatic test site as per protocol (for events dated 17 April-11 May)?

Outcome: Pre-event LFD test completed at an asymptomatic test site vs at home (binary variable). Estimates will be presented with 95% confidence intervals.

Analyses: Adherence to the study protocol around ATS will be assessed for the earlier events will be described using summary tables, as demonstrated in Table 4, per event and pooled. Differences between baseline characteristics in people who complete a pre-event LFD test at an ATS, and those who complete the test at home, will be reported.

Data will be stratified by sociodemographic factors, vaccination status, previous infection status, indoor/outdoor event, seated/standing/mixed event, mean crowd density, relative environmental transmission risk and method of accessing PCR testing. For tables based on single events, the following univariate tests will be used: if normally distributed continuous data, a one-way analysis of variance (ANOVA) test will be used to assess for differences. If continuous data is not normally distributed, non-parametric Kruskal-Wallis tests will be used. If the data is categorical, the Freeman-Halton exact test will be used to assess for differences. For the table based on pooled data, the univariate statistical tests will take into account the correlation within event (cluster design).

Table 4. Summary table describing pre-event LFD site. One table will be presented per event and an additional table presented for data pooled across all events.

		Total	**Pre-event LFD site**			***p*-value**
			Home	**Asymptomatic testing site**	**Unknown**	
Sex	Males	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	x.xx*
	Females	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Unknown/NA	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
Age	Mean (SD)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	x.xx**
	Range (min, max)	xx.x-xx.x	xx.x-xx.x	xx.x-xx.x	xx.x-xx.x	
Ethnicity	Black, African, Black British, or Caribbean	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	x.xx*
	Asian or Asian British	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Another ethnic group	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Mixed or multiple ethnic group	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	White	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Prefer not to say	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Unknown/NA	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
Index of Multiple Deprivation decile	1	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	x.xx*
	2	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	

		Total	Pre-event LFD site			*p*-value
			Home	Asymptomatic testing site	Unknown	
	3	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	4	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	5	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	6	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	7	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	8	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	9	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	10	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
COVID-19 test positive in 2021	Yes	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	x.xx*
	No	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Unknown/NA	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
Vaccinated	1st dose	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	x.xx*
	2nd dose	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	None	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Unknown/NA	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
Vaccinated or Previous Covid	Yes	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	x.xx*
	No	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Unknown/NA	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
Event type***	Indoor – poorly ventilated	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	x.xx*
	Indoor – mechanically ventilated to COVID guidance standards	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Indoor – naturally ventilated/low ventilation	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Indoor – naturally ventilated/high ventilation	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Outdoors – sheltered	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	

		Total	Pre-event LFD site			*p*-value
			Home	Asymptomatic testing site	Unknown	
	Outdoors	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
Audience***	Seated	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	x.xx*
	Mixed	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Standing	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
Crowd density***	Mean (sd)	xx	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	x.xx**
	Range (min, max)	xx	xx.x-xx.x	xx.x-xx.x	xx.x-xx.x	
Method of accessing PCR testing ***	Self-request	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	x.xx*
	Pick-up from pre-event LFD asymptomatic testing site	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Automatically posted	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	

For single events: *Freeman-Halton exact test, **ANOVA/Kruskall-Wallis test. For data with pooled data statistical tests will take into account the clustering nature of the data. ***for inclusion in pooled table only

1b. What proportion of those who attend the event return a completed pre-event home test PCR kit?

Outcome: Pre-event PCR returned vs not returned/unknown (binary variable). Not returned and unknown have been grouped together because it may not be possible to distinguish between the two due to data quality issues (missing NHS numbers, or incomplete or inaccurate self-reported data fields). Estimates will be presented with 95% confidence intervals.

Analyses: Summary descriptive statistics will be presented in tables as demonstrated in Table 5, per event and pooled. Differences between baseline characteristics in people who return a pre-event PCR, and those for whom there has been no pre-event PCR identified, will be reported.

Data will be stratified by sociodemographic factors, vaccination status, previous infection status, indoor/outdoor event, seated/standing/mixed event, mean crowd density, relative environmental transmission risk and method of accessing PCR testing. For tables based on single events, the following univariate tests will be used: if normally distributed continuous data, an independent t-test will be used to assess for differences. If continuous data is not normally distributed, non-parametric Mann-Whitney U tests will be used. If the data is categorical, Chi-squared or the Fisher's exact test will be used to assess for differences. For the table based on pooled data, the univariate statistical tests will take into account the correlation within event (cluster design).

A generalised linear mixed effects model that takes into account the clustered nature of the data (grouped by event) will be used to identify individual characteristics/factors that are associated with PCR test uptake rate. Event-level characteristics may be accounted for in the model assuming there are sufficient numbers of clusters to allow meaningful interpretation.

Table 5. Summary table describing pre-event PCR response rate. One table will be presented per event and an additional table presented for data pooled across all events.

		Total	Pre-event home PCR test kit returned		*p*-value
			Yes	No/Unknown	
Sex	Males	xx	xx (xx.x%)	xx (xx.x%)	x.xx*
	Females	xx	xx (xx.x%)	xx (xx.x%)	
	Unknown/NA	xx	xx (xx.x%)	xx (xx.x%)	
Age	Mean (sd)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	x.xx**
	Range (min, max)	xx.x-xx.x	xx.x-xx.x	xx.x-xx.x	
Ethnicity	Black, African, Black British, or Caribbean	xx	xx (xx.x%)	xx (xx.x%)	x.xx*
	Asian or Asian British	xx	xx (xx.x%)	xx (xx.x%)	
	Another ethnic group	xx	xx (xx.x%)	xx (xx.x%)	
	Mixed or multiple ethnic group	xx	xx (xx.x%)	xx (xx.x%)	
	White	xx	xx (xx.x%)	xx (xx.x%)	
	Prefer not to say	xx	xx (xx.x%)	xx (xx.x%)	
	Unknown/NA	xx	xx (xx.x%)	xx (xx.x%)	
Index of Multiple Deprivation decile	1	xx	xx (xx.x%)	xx (xx.x%)	x.xx*
	2	xx	xx (xx.x%)	xx (xx.x%)	
	3	xx	xx (xx.x%)	xx (xx.x%)	
	4	xx	xx (xx.x%)	xx (xx.x%)	
	5	xx	xx (xx.x%)	xx (xx.x%)	
	6	xx	xx (xx.x%)	xx (xx.x%)	
	7	xx	xx (xx.x%)	xx (xx.x%)	
	8	xx	xx (xx.x%)	xx (xx.x%)	
	9	xx	xx (xx.x%)	xx (xx.x%)	
	10	xx	xx (xx.x%)	xx (xx.x%)	
COVID-19 test positive in 2021	Yes	xx	xx (xx.x%)	xx (xx.x%)	x.xx*
	No	xx	xx (xx.x%)	xx (xx.x%)	

		Total	Pre-event home PCR test kit returned		*p*-value
			Yes	No/Unknown	
	Unknown/NA	xx	xx (xx.x%)	xx (xx.x%)	
Vaccinated	1st dose	xx	xx (xx.x%)	xx (xx.x%)	x.xx*
	2nd dose	xx	xx (xx.x%)	xx (xx.x%)	
	None	xx	xx (xx.x%)	xx (xx.x%)	
	Unknown/NA	xx	xx (xx.x%)	xx (xx.x%)	
Vaccinated or Previous Covid	Yes	xx	xx (xx.x%)	xx (xx.x%)	x.xx*
	No	xx	xx (xx.x%)	xx (xx.x%)	
	Unknown/NA	xx	xx (xx.x%)	xx (xx.x%)	
Event type***	Indoor – poorly ventilated	xx	xx (xx.x%)	xx (xx.x%)	x.xx*
	Indoor – mechanically ventilated to COVID guidance standards	xx	xx (xx.x%)	xx (xx.x%)	
	Indoor – naturally ventilated/low ventilation	xx	xx (xx.x%)	xx (xx.x%)	
	Indoor – naturally ventilated/high ventilation	xx	xx (xx.x%)	xx (xx.x%)	
	Outdoors – sheltered	xx	xx (xx.x%)	xx (xx.x%)	
	Outdoors	xx	xx (xx.x%)	xx (xx.x%)	
Audience***	Seated	xx	xx (xx.x%)	xx (xx.x%)	x.xx*
	Standing	xx	xx (xx.x%)	xx (xx.x%)	
	Mixed	xx	xx (xx.x%)	xx (xx.x%)	
Crowd density***	Mean (sd)	xx	xx.x (xx.x)	xx.x (xx.x)	x.xx**
	Range (min, max)	xx	xx.x-xx.x	xx.x-xx.x	
Method of accessing PCR testing ***	Self-request	xx	xx (xx.x%)	xx (xx.x%)	x.xx*
	Pick-up from pre-event LFD asymptomatic testing site	xx	xx (xx.x%)	xx (xx.x%)	
	Automatically posted	xx	xx (xx.x%)	xx (xx.x%)	

For single events: *Chi-squared test/Fisher's exact test, **independent t-test/Mann-whitney U test. For table with pooled data, statistical tests will take into account the clustering nature of the data. **For inclusion in pooled table only

1c. What proportion of those who attend an event return a completed post-event home test PCR kit?

Outcome: Post-event PCR returned vs not returned/unknown (binary variable). Estimates will be presented with 95% confidence intervals.
Analyses: As per 1b.

1d. What proportion of those who attend the event return both completed pre- and post-event home test PCR kits?

Outcome: Pre- and post-event PCR returned vs only one returned/neither returned/unknown (binary variable).
Analyses: As per 1b.

2a. What proportion of those who attend the events receive a PCR positive test result indicative of them attending the event already infected (i.e. index cases)?

Outcome: Pre-event PCR result (positive/negative/void).
Analyses: Pre-event PCR positive rates will be presented in tables as demonstrated in Table 6, by event and pooled across all events. Estimates will be presented with 95% confidence intervals.

Table 6. Pre-event PCR results.

Event (date)	Pre-event PCR Result	Summary	Event (date)	Pre-event PCR Result	Summary
ACC Business Event (28/04/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)	Snooker World Championships (21/04/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)
Brit Awards (11/05/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)	Snooker World Championships (22/04/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)
Carabao Cup (25/04/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)	Snooker World Championships (23/04/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)
Circus Nightclub (30/04/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)	Snooker World Championships (24/04/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)
Circus Nightclub (01/05/21)	n n unknown/missing positive, n (%) negative, n (%)	xx xx xx (xx.x%)	Snooker World Championships (25/04/21)	n n unknown/missing positive, n (%) negative, n (%)	xx xx xx (xx.x%)

Event (date)	Pre-event PCR Result	Summary	Event (date)	Pre-event PCR Result	Summary
	void, n (%)	xx (xx.x%)		void, n (%)	xx (xx.x%)
FA Cup Final (15/05/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)	Snooker World Championships (26/04/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)
FA Cup Semi-final (18/04/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)	Snooker World Championships (27/04/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)
Live National Outdoor Music Event (02/05/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)	Snooker World Championships (28/04/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)
Mass Participation Run (15/05/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)	Snooker World Championships (29/04/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)
Snooker World Championships (17/04/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)	Snooker World Championships (30/04/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)
Snooker World Championships (18/04/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)	Snooker World Championships (01/05/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)

Event (date)	Pre-event PCR Result	Summary	Event (date)	Pre-event PCR Result	Summary
Snooker World Championships (19/04/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)	Snooker World Championships (02/05/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)
Snooker World Championships (20/04/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)	Snooker World Championships (03/05/21)	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)
-	-	-	Pooled	n n unknown/missing positive, n (%) negative, n (%) void, n (%)	xx xx xx (xx.x%) xx (xx.x%) xx (xx.x%)

Summary descriptive tables of baseline characteristics by pre-event PCR result (positive, negative, void or unknown) will be reported as demonstrated in Table 7. In particular, data will be stratified by sociodemographic factors, vaccination status, previous infection status, indoor/outdoor event, seated/standing/mixed event, mean crowd density and relative environmental transmission risk. The count and proportion of those for which an NHS number has not been matched and where fuzzy matching on demographic details from ticketing to NHS records has failed – i.e. PCR result unknown – will be reported.

Table 7. Summary table describing demographics stratified by pre-event PCR result. One table will be presented per event and an additional table presented for data pooled across all events.

		Total	Pre-event home PCR test result				*p*-value
			Positive	Negative	Void	Unknown	
Sex	Males	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	x.xx*
	Females	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Unknown/NA	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
Age	Mean (sd)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	x.xx**
	Range (min, max)	xx.x-xx.x	xx.x-xx.x	xx.x-xx.x	xx.x-xx.x	xx.x-xx.x	
Ethnicity	Black, African, Black British, or Caribbean	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	x.xx*
	Asian or Asian British	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	

		Total	Pre-event home PCR test result				*p*-value
			Positive	Negative	Void	Unknown	
	Another ethnic group	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Mixed or multiple ethnic group	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	White	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Prefer not to say	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Unknown/NA	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
Index of Multiple Deprivation decile	1	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	x.xx*
	2	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	3	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	4	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	5	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	6	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	7	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	8	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	9	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	10	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
COVID-19 test positive in 2021	Yes	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	x.xx*
	No	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Unknown/NA	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
Vaccinated	1st dose	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	x.xx*
	2nd dose	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	None	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Unknown/NA	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
Vaccinated or Previous Covid	Yes	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	x.xx*
	No	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Unknown/NA	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	

		Total	Pre-event home PCR test result				*p*-value
			Positive	Negative	Void	Unknown	
Event type***	Indoor – poorly ventilated	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	x.xx*
	Indoor – mechanically ventilated to COVID guidance standards	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Indoor – naturally ventilated/low ventilation	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Indoor – naturally ventilated/high ventilation	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Outdoors – sheltered	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Outdoors	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
Audience ***	Seated	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	x.xx*
	Standing	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
	Mixed	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)	
Crowd density***	Mean (sd)	xx	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	x.xx**
	Range (min, max)	xx	xx.x-xx.x	xx.x-xx.x	xx.x-xx.x	xx.x-xx.x	

For single events: *Freeman-Halton exact test, **ANOVA/Kruskal-Wallis test. For data with pooled data statistical tests will take into account the clustering nature of the data. ***for inclusion in pooled table only

For the Mass Participation Run event, where eventgoers opted to have their pre-event LFD test either at home or at an asymptomatic test site, differences in socio-demographic characteristics between the two groups will be assessed. Baseline characteristics will be presented as demonstrated in Table 8. Generalised linear modelling will be applied to identify individual characteristics/factors that are associated with LFD testing choice.

Table 8. Summary table comparing the baseline characteristics of Mass Participation Run eventgoers who underwent pre-event LFD testing at home vs. at an asymptomatic test site.

		Total	Site of pre-event LFD test		
			Home	Asymptomatic test centre	Unknown
Sex	Males	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
	Females	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
	Unknown/NA	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
Age	Mean (sd)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)	xx.x (xx.x)
	Range (min, max)	xx.x-xx.x	xx.x-xx.x	xx.x-xx.x	xx.x-xx.x
Ethnicity	Black, African, Black British, or Caribbean	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)

		Total	Site of pre-event LFD test		
			Home	Asymptomatic test centre	Unknown
	Asian or Asian British	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
	Another ethnic group	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
	Mixed or multiple ethnic group	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
	White	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
	Prefer not to say	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
	Unknown/NA	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
Index of Multiple Deprivation decile	1	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
	2	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
	3	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
	4	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
	5	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
	6	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
	7	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
	8	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
	9	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
	10	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
COVID-19 test positive in 2021	Yes	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
	No	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
	Unknown/NA	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
Vaccinated	1st dose	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
	2nd dose	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
	None	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
	Unknown/NA	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
Vaccinated or Previous Covid	Yes	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)
	No	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)

		Total	Site of pre-event LFD test		
			Home	Asymptomatic test centre	Unknown
	Unknown/NA	xx	xx (xx.x%)	xx (xx.x%)	xx (xx.x%)

The relative risk of a receiving a positive/negative/void/unknown result depending on where the LFD test was taken (home or asymptomatic site) will be calculated for the Mass Participation Run and the FA Cup Final events, as demonstrated in Table 8. Note that this study does not have sufficient power to test the hypothesis of non-inferiority.

Table 9. Difference in pre-event PCR test result by method of pre-event LFD testing. One table will be presented per event and an additional table presented for data pooled across both events.

	Home LFD testing no./total no. (%)	**Asymptomatic LFD testing** no./total no. (%)	**Relative risk** (95% CI)	***p*-value**
No. of participants	xx	xx	x.xx (x.xx-x.xx)	x.xx*
Positive	xx/xx (xx.x%)	xx/xx (xx.x%)	x.xx (x.xx-x.xx)	x.xx*
Negative	xx/xx (xx.x%)	xx/xx (xx.x%)	x.xx (x.xx-x.xx)	
Void	xx/xx (xx.x%)	xx/xx (xx.x%)	x.xx (x.xx-x.xx)	
Unknown	xx/xx (xx.x%)	xx/xx (xx.x%)	x.xx (x.xx-x.xx)	

*chi-squared test/Fisher's exact test

2b. What proportion of those who attend the events receive a post-event PCR positive test result and is suspected to be a secondary case (e.g. pre-event PCR negative test result)?

Outcome: Post-event PCR result (positive/negative/void).
Analyses: As per 2a.

3a. Using whole genome sequencing of positive home test PCR kits, how many eventgoer cases detected on post-event home test PCR are linked by their genomic sequencing results (i.e. secondary cases)?

Primary outcome: Characterisation of SARS-CoV-2 viruses from post-event PCR test positives.
Analysis of primary outcome: Descriptive – counts and proportions – as demonstrated in Table 9.
Secondary outcomes: Nil.
Analysis of secondary outcomes: Nil.

3b. Can genomic sequencing of pre-event PCR positive tests identify a linked index case?

Primary outcome: Characterisation of SARS-CoV-2 viruses from pre-event PCR test positives.
Analysis of primary outcome: Descriptive – counts and proportions – as demonstrated in Table 9.
Secondary outcomes: Nil.
Analysis of secondary outcomes: Nil.

Table 10. Summary table of genomically-confirmed clusters.

	Pre-event PCR	**Post-event PCR**	**Cluster size**
No. of positive PCR test results	xx (xx.x%)	xx (xx.x%)	
No. of positive PCR test results where genomic sequencing was possible	xx (xx.x%)	xx (xx.x%)	
Strain 1	xx (xx.x%)	xx (xx.x%)	xx
Strain 2	xx (xx.x%)	xx (xx.x%)	xx

	Pre-event PCR	Post-event PCR	Cluster size
Strain 3 etc	xx (xx.x%)	xx (xx.x%)	xx
No of positive PCR test results where genomic sequencing was not possible	xx (xx.x%)	xx (xx.x%)	

5.2. Sensitivity analyses and missing data

Missing data will not be imputed – a complete case analysis approach will be used. Due to the expected low number of positive tests, sub-group analyses will not be performed.

5.3. Statistical software

Statistical analyses will be performed using SPSS version 25 at PHE and R version 4.0.5 at University of Liverpool for the Liverpool events (See Re-Liv-E SAP). Analyses will be checked by a second PHE researcher using Stata.

6. Discussion

The overall aim of the study is to provide evidence on the feasibility of pre-event rapid antigen testing with lateral flow devices (LFD) for SARS-CoV-2 in mitigating the risk of COVID-19 transmission amongst spectators, participants or audiences at cultural and sporting events, instrumented with and pre- and post-event PCR testing. The broad aim has generated several research objectives and questions to be answered. This SAP has been drafted to avoid outcome reporting bias when answering the national research questions, but there are a number of remaining limitations that should be recognised:

- Selection bias:
    - Self-selection bias: people who choose to attend events are likely to be more risk-tolerant, and may have increased risk of exposures to COVID-19.
    - Attrition bias: if free ticketing to certain events.
    - Sampling bias: some events are 'invite-only' to certain groups – e.g. NHS key workers at the Brits.
    - Susceptibility bias: current guidance states that people who have received a positive test in the previous 90 days should not participate in further COVID-19 testing unless new symptoms develop – therefore these people may be less likely to participate in events.
- Non-response bias – people may choose not to return a PCR test – differences between those who do and do not return PCR tests will be examined.
- Falsified results/gaming – there is a possibility that some eventgoers may falsify either the proof of their negative LFD result, or purposely use improper technique to ensure a negative LFD result. This risk will not be evaluated as part of this study.
- Residual confounding – there are likely to be a number of additional confounding factors on which data was not collected as part of this study.
- Data quality challenges – matching eventgoers to their test results and other datasets requires the event organisers to share an attendee list comprised of high-quality data fields (full name, gender, date of birth, full address including postcode) with PHE. These fields are self-reported to the event organiser by the eventgoer – the quality, accuracy and completeness of these date fields cannot be ensured by PHE.
- PCR test return challenges – return of PCR tests is not incentivised or mandated for Phase 1 ERP participants. No detection of SARS-CoV-2 by PCR test would not mean that there were no index cases at the event and/or no transmission occurred at the event – index or

secondary cases may be present in those who do not return PCR tests. Even with 100% PCR return, the tests are unsupervised meaning some uncertainty of sample quality remains.

Finally, this work is likely to generate further research questions and hypotheses that could be tested going forward.

7. Conclusion

This SAP presents the principles of analysis of the LFD testing, PCR testing and SARS-CoV-2 viral genomic sequencing data from the ERP repeated cross-sectional study, to inform discussions around transmission risk mitigation for COVID-19. We hope that the results of this study will be as transparent and robust as possible, to inform the design and development of future phases of the ERP, and ultimately to feed into the Government's Roadmap for COVID-19 recovery.