

Investigation of high risk exposure settings for COVID-19 in England

This paper describes a rapid case control study implemented by Public Health England working with NHS Test and Trace to provide insights on settings and activities associated with risk of COVID-19 infection.

The study was implemented quickly to inform public health action and policy and it is based on the comparison on information obtained from cases through the contact tracing process and general population controls obtained through a market research company.

The study has now been run monthly for the past 3 months. This report contains the results of the third iteration with some information on the results of the previous two in an appendix.

The study has limitations which are described in the report. The impact of these has been assessed where possible.

The study shows that certain occupational groups are associated with increased odds of being COVID-19 case. It also shows that engaging in entertainment activities is associated with higher odds of being a COVID-19 case.

The results presented in this study are consistent with evidence from other studies and the previous iterations of this case-control study.

This information should be used together with other insights on settings and activities associated with transmission of COVID-19 to inform control measures in England.

Work is underway in collaboration with academic partners to strengthen the study.

Public Health England and NHS T&T

11 November 2020

Summary of results from an epidemiological investigation

1 Summary

This study reports preliminary results of the third epidemiological investigation into high-risk exposure settings for COVID-19 in England using data collected through enhanced contact tracing and a market research panel.

The study showed that there was strong statistical evidence that working in warehouse settings and construction, education and hospitality, as well as health and social care was associated with increased odds of being COVID-19 case. Additionally, there was strong statistical evidence that engaging in entertainment activities was associated increased odds of disease. There was no clear evidence of an increased odds of disease associated with visiting a closed space or crowded area. When adjusting for misclassification arising from imperfect sensitivity and specificity of RT-PCR test method, exposures associated in the main multivariable analysis remained associated with increased odds of being a COVID-19 case.

These findings should be viewed in the context of the limitations of the study. However, the results presented in this study show strong consistency with the existing evidence from other studies, and also from the previous iterations of this case-control study. As such, the study provides evidence that can inform the COVID-19 control measures in England.

2 Background

As part of the response to the COVID-19 pandemic in England, diagnostic laboratories have a statutory duty to notify Public Health England (PHE) of all confirmed cases of COVID-19. Positive SARS-CoV-2 test results are reported by PHE to the NHS Test and Trace service. The role of NHS Test and Trace is to ensure that recent close contacts of all positive cases are identified, and that relevant public health advice and information can be provided. NHS Test and Trace service will directly contact a case via text, email or telephone to request completion of a questionnaire, which may be completed electronically or by telephone interview with a contact tracer. NHS Test and Trace seeks to identify all contacts considered to have had a risk of exposure. Early results from mass contact tracing efforts show that it can be an effective tool in bringing down transmission rates in the community, leading to a desired epidemiological change (Kendall et al., 2020). Since 10th of August 2020, the questionnaires have been enhanced to include detailed information about exposures in the seven days prior to symptom onset for all cases reported; referred to as 'Enhanced Contact Tracing'.

Enhanced contact tracing can provide data on the potential exposure settings where the transmission of SARS-CoV-2 occurs. Understanding the role of varied community environments in the transmission of SARS-CoV-2 is critical in enabling an effective public health response. The data collected via enhanced contact tracing offers a unique opportunity to understand the risk associated with different exposure settings. Evidence towards high-risk exposures will help guide public health action and may, in effect, help to reduce the transmission across a variety of community settings.

There are numerous studies demonstrating the importance of household transmission for the spread of COVID-19 (Lei et al., 2020). However, there is currently limited evidence on the importance of individual community settings beyond the household on COVID-19 transmission. A study by the Centers for Disease Control and Prevention (CDC) in the United States found that COVID-19 cases were more likely to have reported visiting a restaurant or bar/coffee shop than non-cases, although this study did not distinguish between dining indoors or outdoors and had a limited and potentially biased sample population (Fisher et al., 2020).

A review of published outbreak clusters conducted in April 2020 also identified numerous clusters associated with restaurants and cafes, although few clusters were found in schools and hotels (Leclerc et al., 2020). This research, however, was limited to studies published during the first phase of the U.K. pandemic and therefore does not reflect the impact of non-pharmaceutical public health measures implemented more recently. The risk on healthcare workers has been reported in the recent prospective cohort study in the United Kingdom (Nguyen et al., 2020). Another study also reported outbreaks and infections of SARS-CoV-2 being associated with educational settings in the United Kingdom, drawing attention particularly to controlling transmission among students and in staff members (Ismail et al., 2020). A recent technical report published by the European Centers of Disease Control (ECDC) further highlights that most of the reported local outbreaks and clusters of COVID-19 are associated with occupational and healthcare settings. The link to occupational settings is reported particularly true for sectors considered as essential (ECDC, 2020).

Overall, the evidence on high-risk social activity settings for COVID-19 is sparse. Improved understanding on high-risk settings provides critical evidence to guide the public health policies and control measures. This study aims to address some of these gaps through a retrospective case-control study, which aims to identify settings associated with increased risk for COVID-19.

The study has been repeated three times, the first study taking place early September, the second late September and the most recent study, described in this report, late October. The results from first two studies have highlighted certain exposures being associated with increased odds of illness. The multivariable analyses from the first study provided strong statistical evidence that working in hospitality, healthcare and social care were associated with increased odds of being a COVID-19 case. The results also showed some evidence that cases were more likely to work in warehouse and close contact services than controls. The first study gave strong signal that entertainment activities were associated with increased odds of being a COVID-19 case (See Appendix 4 Table 1 for full

results). The second study showed strong evidence of an association between being a COVID-19 case and working in hospitality or in social care. The results from the second study also showed some evidence that cases were more likely to work in a warehouse or in healthcare setting than controls. The results also showed strong evidence that going to pub or bar and engaging in other entertainment activities were associated with increased odds of illness (see Appendix 4 Table 2 for full results).

2 Methods

2.1 Study design and population

The study is a retrospective, frequency-matched case-control study. The present report is the third iteration of this study and the summaries of previous results are provided in Appendix 3. The study population consists of adults over 18 years old who reside in England and Scotland. 4,223 individuals were recruited – 2,000 cases and 2,223 controls. Cases were confirmed cases of COVID-19 that had completed the NHS Test and Trace Enhanced Contact tracing questionnaires. Cases had provided information either through a digital route (self-completed) or been interviewed by phone. The controls were members of the general public who had not tested positive for SARS-CoV-2 and were asymptomatic and were not household contacts of confirmed cases. The controls were recruited through Market Research Panel (MRP) company pool of volunteers.

Cases were defined as any person above 18 years of age, with a laboratory confirmed SARS-CoV-2 positive result who had also completed contact tracing. Based on pre-defined exclusion criteria, we did not include those with undefined sample dates or incomplete NHS Test and Trace profiles, or those who were identified being a household contact of a confirmed case of COVID-19 prior their onset of illness.

Controls were defined as any person above 18 years of age residing in England, and a volunteer for MRP. We excluded potential controls who had had positive SARS-CoV-2 test, were a household contact of a case or had a history of symptoms of COVID-19 during 7 days prior to completion of the questionnaire.

In addition to the specific exclusion criteria outlined above, we also excluded any cases or controls whose responses showed evidence of excessive responses. Excessive responses were defined by introducing plausible cut-offs for the number of activities reported for the 7 day exposure period (for further details please see Appendix 3).

2.1.3. Sampling

Cases and controls were frequency matched by age and region. Random sampling was used to extract the records for cases from NHS T&T databases.

Regarding the selection of controls, the MRP service made the control questionnaire available to their members on members' dashboards. Members were shown basic information about the survey (e.g., completion time, expected pay) that did not include any information about the survey content (Pureprofile, 2017). Members could then choose whether to engage with the survey. The completion rate for the present round of controls was 24%, so 24% of the MRP service members who were shown our survey on their dashboards completed the questionnaire.

Based on publicly available data, MRP services are able to target surveys at individuals with specific characteristics (e.g., age, gender, location; Pureprofile, 2017), thus providing a frequency-matched sample of controls. Presently, controls were frequency matched by 3 age groups (18-29 years, 30-49, 50 and above) and by geography (London, South England, North England, Midlands).

2.1.4. Time period of recruitment and exposure period

The surveys for controls were deployed on 22nd October, and the deployment was over on 25th October. This meant that the beginning of the exposure period for controls ranged between 15th to 17th October.

Cases were extracted by the sample date of their positive test, using a timeframe (23rd to 25th October) that was matched to as closely to that of controls as possible. The exposure period for cases is less well defined, as they are asked to recall their exposures 7 days prior the onset of their illness. As the onset date is less accurately recorded, sample dates were used as a proxy. Therefore, it can be assumed that the beginning of the exposure period for cases ranged between 13th to 15th October.

2.3. Analyses

2.3.1 Descriptive analysis

Cases and controls were described according to their demographics (age, sex, ethnicity, region of residence and index of multiple deprivation (IMD)).

2.3.2 Univariable and multivariable analysis

Univariable (logistic regression) analyses were undertaken to estimate crude Odds Ratios (cORs) as measures of association between exposures (work and leisure activities) and binary outcome (COVID-19 case yes/no). Confidence intervals around these estimates and p-values were calculated. We also calculated the population attributable fraction (PAF) for specific exposures (see Appendix 2 for formulas). PAF estimates how much of the incidence in the population can be attributed to certain risk factors.

We used Firth correction (penalised regression method; Firth, 1993) in our multivariable analyses. This type of model provides more meaningful results in situations of low prevalence of exposure, which is the case for many of the exposures in the study. The phenomenon of separation is observed in the

fitting process of a logistic model if the likelihood converges while at least one parameter estimate diverges to \pm infinity. Separation primarily occurs in small samples with several unbalanced and highly predictive risk factors. A procedure by Firth originally developed to reduce the bias of maximum likelihood estimates is shown to provide an ideal solution to separation. It produces finite parameter estimates by means of penalized maximum likelihood estimation (Heinze et al. 2002).

Penalised regression methods were used to estimate the odds ratios for exposures adjusted for effects of other exposures and demographic variables. Main exposures (work and leisure activities) were all included in the model. Specific variables of interest were considered to be included in the model if there was evidence of an association (p-value <0.2, increased OR, and proportion of cases exposed was at least 0.4%). Demographic variables were included as potential confounders and were considered to be excluded in case of multicollinearity.

2.3.3 Sensitivity analysis – accounting for misclassification of outcome

To account for potential misclassification due to imperfect sensitivity and specificity of RT-PCR in detecting SARS-CoV-2, we conducted sensitivity analyses which were based on an iterative expectation-maximization (EM) algorithm (Magder et al., 1997). Sensitivity of RT-PCR in detecting SARS-CoV-2 reported in the literature (Arevalo-Rodriguez et al., 2020, Böger et al., 2020; cf. Watson et al., 2020; Yu et al., 2020), ranges from 71% to 98%. Specificity reported in the literature above ranges between 90% and 100%. Due to the estimated positivity being around 1% (ONS, 2020), it is unlikely that the imperfect specificity and sensitivity will lead to substantial misclassification of the case/non-case status. However, to explore the impact for this potential misclassification, we will test misclassification that could arise from sensitivity and specificity of 90%, respectively. The outcomes of these analyses will be compared to that of the main MVA. The sensitivity in the context our analyses refers to the proportion of those tested positive that are true cases.

2.3.4 Sensitivity analysis – accounting for the nature of the activity setting

The survey included specific questions related to the context in which working and leisure activities occurred. For each exposure it was ascertained whether the exposure occurred in a crowded space or not. Crowded in these questions was described as: “*difficulty keeping two meters apart from others*”. We also performed a similar analysis accounting for whether the activity took place in a closed space, a closed space being described as “*indoor room, hall or office, inside a vehicle*”. Using a penalized regression model, we explored if activities taking place in these types of settings would modify our estimate of the measures of effect in any direction. We also explored if being in a closed or crowded place would show a consistent trend increasing the odds ratio of becoming a case.

2.3.5 Sensitivity analysis – accounting for the mode of questionnaire completion (online/telephone)

While all controls completed the questionnaire online on the MRP service's website, not all cases completed the questionnaire online. Cases who failed to complete the questionnaire online were escalated to be followed up by contact tracers or HPTs.

Using multinomial models we explored the effect of mode of questionnaire completion on the associations of exposures with the outcome of being a COVID-19 case.

3 Results

3.1. Study population

The study recruited 4,223 participants, including 2,000 cases and 2,223 controls. The median number of exposures reported by cases was 2 (range 1-20), while for controls the median number of responses was 5 (range 1- 68). Based on the pre-defined exclusion criteria, we excluded 149 controls due to excessive responses (and 0 cases). This left us with 2,000 cases and 2,074 controls.

3.2. Demographics

The sex distribution was very similar between cases and controls. A higher proportion of cases than controls were in the age group 48-57 years. The median age and age range again were nearly identical. More cases than controls lived in the East of England, South East, South West, and Yorkshire and Humber. More cases than controls identified their ethnicity as "other" and there was a higher proportion of controls identifying as white. While information on ethnicity was available for all controls, it was missing for 10% of cases. The postcodes of cases were more likely to be in least deprived areas compared to controls. IMD distribution otherwise was similar, but information on deprivation was missing for 11% of controls.

Table 1. Demographic distribution of cases and controls

		Cases (n = 2000)		Controls (n = 2074)	
		n	p	n	p
Sex	Male	891	45	889	43
	Female	1069	53	1182	57
	Other sex	0	0	3	0
	Missing sex	40	2	0	0
Age group (years)	18-27	490	24	505	24
	28-37	458	23	471	23
	38-47	368	18	411	20
	48-57	408	20	315	15
	58+	276	14	372	18
	Missing age	0	0	0	0
Median age (range)	39 (18-92)		39 (18-98)		
Location	East of England	119	6	40	2
	East Midlands	226	11	233	11
	London	183	9	210	10
	North East	111	6	286	14
	North West	507	25	740	36
	South East	163	8	110	5
	South West	137	7	62	3
	West Midlands	226	11	293	14
	Yorkshire and Humber	328	16	100	5
	Missing location	0	0	0	0
Ethnicity	White	1426	71	1804	87
	Mixed	50	2	60	3
	Asian	168	8	126	6
	Black	32	2	63	3
	Other ethnicity	132	7	21	1
	Missing ethnicity	192	10	0	0
Index of Multiple Deprivation	1 (most deprived)	510	26	557	27
	2	410	20	420	20
	3	377	19	326	16
	4	380	19	302	15
	5 (least deprived)	323	16	248	12
	Missing IMD	0	0	221	11

3.2. Univariable analysis – grouped exposures

Table 2 describes the results from univariable analysis with crude odds ratios (cORs). There was evidence that working in warehouse settings, hospitality, healthcare, social care, construction, education, and close contact services were crudely associated with being a COVID-19 case. There was also evidence that working in emergency services and transport as well as engaging

entertainment activities was crudely associated with being a COVID-19 case. There was no evidence to suggest that cases were more likely to have worked in retail, food production, or travel.

Table 2. Univariable analysis of higher level exposures

Exposure	Cases			Controls			Odds Ratio	95% CI	P-value
	Total	Exposed	% exposed	Total	Exposed	% exposed			
Work and education									
Warehouse ^a	2000	53	2.65	2074	4	0.19	14.09	[5.17-53.67]	< 0.001
Close contact ^b	2000	38	1.9	2074	8	0.39	5	[2.29-12.44]	< 0.001
Military ^g	2000	10	0.5	2074	3	0.14	3.47	[0.89-19.64]	0.044
Healthcare	2000	197	9.85	2074	75	3.62	2.91	[2.20-3.88]	< 0.001
Construction ^c	2000	154	7.7	2074	59	2.84	2.85	[2.08-3.94]	< 0.001
Emergency services	2000	25	1.25	2074	10	0.48	2.61	[1.21-6.11]	0.008
Social care	2000	52	2.6	2074	25	1.21	2.19	[1.33-3.69]	0.001
Hospitality ^d	2000	86	4.3	2074	47	2.27	1.94	[1.33-2.84]	< 0.001
Arts and recreation ^h	2000	30	1.5	2074	17	0.82	1.84	[0.98-3.57]	0.042
Transport ^f	2000	59	2.95	2074	34	1.64	1.82	[1.17-2.88]	0.005
Education ^e	2000	290	14.5	2074	198	9.55	1.61	[1.32-1.96]	< 0.001
Food production	2000	17	0.85	2074	17	0.82	1.04	[0.50-2.17]	0.915
Retail	2000	73	3.65	2074	74	3.57	1.02	[0.73-1.44]	0.888
Work related travel ⁱ	2000	13	0.65	2074	36	1.74	0.37	[0.18-0.72]	0.001
Leisure activities									
Entertainment	2000	60	3	2074	43	2.07	1.46	[0.97-2.22]	0.06
Sports events ^j	2000	21	1.05	2074	24	1.16	0.91	[0.48-1.71]	0.743
Eating out ^l	2000	312	15.6	2074	475	22.9	0.62	[0.53-0.73]	< 0.001
Personal exercise ^k	2000	150	7.5	2074	272	13.11	0.54	[0.43-0.67]	< 0.001
Private events ⁿ	2000	6	0.3	2074	16	0.77	0.39	[0.12-1.04]	0.04
Travel	2000	26	1.3	2074	86	4.15	0.3	[0.19-0.48]	< 0.001
Community and charity events ^m	2000	6	0.3	2074	25	1.21	0.25	[0.08-0.62]	0.001
Visiting family ^q	2000	116	5.8	2074	438	21.12	0.23	[0.18-0.29]	< 0.001
Personal care ^p	2000	30	1.5	2074	143	6.89	0.21	[0.13-0.31]	< 0.001
Public events, mass gatherings ^o	2000	12	0.6	2074	61	2.94	0.2	[0.10-0.38]	< 0.001
Visiting healthcare ^s	2000	46	2.3	2074	270	13.02	0.16	[0.11-0.22]	< 0.001
Shopping ^t	2000	475	23.75	2074	1455	70.15	0.13	[0.11-0.15]	< 0.001

Note: Working in immigration was included in the model, but removed from the table due to low numbers of participants exposed (0 cases, 1 control).

a. Working in warehouse settings –warehouse, haulage, distribution etc

- b. Working in close contact services – barbers, hairdressers, nail salons, tattoo studios and tanning salons, and any other services which require close contact
- c. Working in manufacturing and construction – textile, electronics, cars
- d. Working in hospitality – working in restaurants, food and drink outlets, lodging etc
- e. Education – working or attending
- f. Working in public transport- including buses, logistics, underground and trains
- g. Working in military – including the Navy, Army and Air Force
- h. Working in arts, or recreation – music, theatre, gyms, cinema, leisure centres
- i. Work related travel – including attending conferences, door-to-door sales, visiting clients
- j. Sports events - including football, cricket, rugby, tennis etc. matches, horse races and other sports events
- k. Personal exercise – e.g. at the gym, swimming, running clubs, casual exercise with friends
- l. Eating out – in different restaurants, cafes, pubs
- m. Charity and community – fundraising, volunteering, bootsale, corporate events, educational classes
- n. Private events – including weddings, funerals, parties, other social gatherings
- o. Mass gatherings – raves, gigs, festivals, protests and live music events
- p. Personal care - Visiting nail salons, hairdressers, barbers, tanning and tattoo studios
- q. Visiting family or friends – indoors or outdoors
- r. Travel - any international and domestic travel including trains and flights
- s. Health care - visiting healthcare services for non-COVID-19 reasons
- t. Shopping at supermarkets, local shops and pharmacies
- u. Other entertainment activities (non-specific)
- v. Going to pub or bar (entertainment, not e.g. eating out)

3.3. Univariable analysis – specific exposures

Univariable analysis for specific risk settings (Tables 3-6) show crude odds ratios for specific activities. Table 3 focuses on entertainment activities, including attending music events, bars and pubs for social gatherings. Table 4 displays any eating out activities. Table 5 examines work settings in relation to warehouses while Table 6 examines work settings in relation to construction (including manufacturing).

Overall, the number of cases exposed to entertainment activities was low (3%) for any entertainment exposure (Table 2). There was evidence of cases being more likely than controls to have been exposed to “other” entertainment activities (cOR 9.94, 95% C.I. 2.39-88.04, $P < 0.001$). Cases who reported taking part on ‘Other Entertainment activities’ reported a very diverse set of exposures linked to these activities. These included activities taking place outdoors, going to theme parks and golf clubs. There was also evidence of cases being more likely than controls to have been eaten out in a café (cOR 2.05, 95% C.I. 1.18-3.66, $P = 0.007$). The estimated PAFs for all activities were low. The estimated PAF was 0.9% for “other” entertainment, and 3.8% for eating out in a café.

Regarding work settings, there was evidence that cases had higher odds than controls to work in “other” warehouse settings (cOR 24.39, 95% C.I. 6.36-207.55, $P < 0.001$) as well as “other” construction/manufacturing settings (cOR 6.73, 95% C.I. 2.59-22.16, $P < 0.001$), in construction labour (cOR 2.62, 95% C.I. 1.12-4.75, $P = 0.013$) and engineering (cOR 2.25, 95% C.I. 1.12-4.75, $P = 0.013$). The types of warehouses that cases reported varied substantially by their nature and size and there was no single commonly reported warehouse setting or company. The cases who reported working in manufacturing and construction also reported very diverse set of employments from self-employed to small or middle-size companies to large companies.

Table 3. Entertainment activities

Exposure	Cases			Controls			Odds Ratio	95% CI	P-value	Population Attributable Fraction*
	Total	Exposed	% exposed	Total	Exposed	% exposed				
Other entertainment ^u	2000	19	0.95	2074	2	0.1	9.94	[2.39-88.04]	< 0.001	0.9%
Visiting a zoo	2000	9	0.45	2074	3	0.14	3.12	[0.78-17.94]	0.072	0.3%
Amusement park	2000	4	0.2	2074	2	0.1	2.08	[0.30-22.97]	0.389	0.1%
Museum visit	2000	4	0.2	2074	2	0.1	2.08	[0.30-22.97]	0.389	0.1%
Day trips	2000	3	0.15	2074	2	0.1	1.56	[0.18-18.65]	0.625	0.05%
Pub or Bar ^v	2000	19	0.95	2074	15	0.72	1.32	[0.63-2.79]	0.426	0.2%
Park	2000	2	0.1	2074	5	0.24	0.41	[0.04-2.53]	0.277	NA
Social club	2000	2	0.1	2074	6	0.29	0.35	[0.03-1.93]	0.172	NA
Cinema visit	2000	2	0.1	2074	23	1.11	0.09	[0.01-0.36]	< 0.001	NA

Note: Activities were either cases or controls had 0 exposed were excluded.

u. Other – cases most commonly reported going to a golf club, taking part in outdoor activities like walks on a beach or sights, going to theme parks

v. Going to a bar or pub for social gatherings, not e.g. eating out

* PAF not reported for exposures with OR less than 1.0

Tables 4a and 4b. Hospitality sector

Working in hospitality

Exposure	Cases			Controls			Odds Ratio	95% C.I.	P-value	Population Attributable Fraction ^a
	Total	Exposed	%	Total	Exposed	%				
Working hospitality - food service	2000	76	3.8	2074	37	1.78	2.17	[1.44-3.33]	<0.001	2.1%

a. Population Attributable Fraction Population attributable fraction (PAF) estimates how much of the incidence in the population can be attributed to certain risk factors. The PAF presented here is a crude estimate. It is not calculated for any exposures with odds ratio less than 1.0

Eating out activities^a

Exposure ^b	Cases			Controls			Odds Ratio	95% C.I.	P-value	Population Attributable Fraction ^a
	Total	Exposed	%	Total	Exposed	%				
Eating out in a café	2000	41	2.05	2074	21	1.01	2.05	[1.18-3.66]	0.007	3.8%
Eating out in a pub or bar	2000	154	7.7	2074	260	12.54	0.58	[0.47-0.72]	<0.001	NA
Eating out in any other restaurant	2000	151	7.55	2074	330	15.91	0.43	[0.35-0.53]	<0.001	NA
Eating out in a market or from mobile food vendor	2000	3	0.15	2074	35	1.69	0.09	[0.02-0.28]	<0.0010	NA

a. Population Attributable Fraction not calculated for exposures which had a crude odds ratio less than 1.0

b. Activities/exposures with no cases exposed not included

Table 5. Warehouse work settings

Exposure	Cases			Controls			Odds Ratio	95% CI	P-value	Population Attributable Fraction
	Total	Exposed	% exposed	Total	Exposed	% exposed				
Other warehouse	2000	46	2.3	2074	2	0.1	24.39	[6.36-207.55]	< 0.001	2.2%
Food warehouse	2000	3	0.15	2074	1	0.05	3.11	[0.25-163.54]	0.3	0.1%
Warehouse haulage	2000	2	0.1	2074	1	0.05	2.08	[0.11-122.47]	0.542	0.05%

Note: Activities were either cases or controls had 0 exposed were excluded. Definitions are not specific and individuals may have interpreted these in different ways

Table 6. Construction (including manufacturing) work settings

Exposure	Cases			Controls			Odds Ratio	95% CI	P-value	Population Attributable Fraction*
	Total	Exposed	% exposed	Total	Exposed	% exposed				
Other construction/manufacturing	2000	32	1.6	2074	5	0.24	6.73	[2.59-22.16]	< 0.001	1.4%
Manufacturing printing	2000	5	0.25	2074	1	0.05	5.2	[0.58-245.81]	0.093	0.2%
Construction labour	2000	35	1.75	2074	14	0.68	2.62	[1.37-5.29]	0.002	1.1%
Construction engineering	2000	28	1.4	2074	13	0.63	2.25	[1.12-4.75]	0.013	0.8%
Manufacturing cars	2000	6	0.3	2074	4	0.19	1.56	[0.37-7.51]	0.49	0.1%
Construction/manufacturing office/administration	2000	34	1.7	2074	27	1.3	1.31	[0.76-2.27]	0.295	0.4%
Manufacturing pharmaceuticals	2000	1	0.05	2074	4	0.19	0.26	[0.01-2.62]	0.193	NA

Note: Activities were either cases or controls had 0 exposed were excluded. Definitions are not specific and individuals may have interpreted these in different ways

* PAF not reported for exposures with OR less than 1.0

3.4. Multivariable analysis

Table 7. Multivariable analysis*

Setting/Activity	Odds Ratio	95% CI	P-value
Work and education			
Warehouse ^a	14.57	[4.39-48.27]	< 0.001
Military ^g	5.37	[0.9-31.93]	0.065
Immigration	3.2	[0.13-82.04]	0.481
Healthcare	3.01	[2.14-4.23]	< 0.001
Construction ^c	2.56	[1.73-3.8]	< 0.001
Social care	2.38	[1.32-4.29]	0.004
Hospitality ^d	2.18	[1.37-3.45]	0.001
Emergency services	2.15	[0.92-5.02]	0.076
Education ^e	1.97	[1.55-2.52]	< 0.001
Close contact ^b	1.97	[0.71-5.51]	0.194
Transport ^f	1.76	[1.04-2.98]	0.034
Arts and recreation ^h	1.72	[0.83-3.55]	0.147
Retail	1.23	[0.82-1.84]	0.314
Food production	0.92	[0.37-2.31]	0.861
Work related travel ⁱ	0.46	[0.22-0.97]	0.04
Leisure activities			
Eating out ^l	0.88	[0.72-1.08]	0.225
Personal exercise ^k	0.87	[0.66-1.14]	0.304
Sports events ^j	0.76	[0.36-1.6]	0.467
Private events ⁿ	0.72	[0.25-2.14]	0.56
Travel ^r	0.43	[0.25-0.74]	0.002
Personal care ^p	0.43	[0.26-0.7]	0.001
Community and charity events ^m	0.4	[0.13-1.25]	0.116
Visiting family ^q	0.4	[0.31-0.51]	< 0.001
Visiting healthcare ^s	0.26	[0.18-0.38]	< 0.001
Public events, mass gatherings ^o	0.19	[0.09-0.43]	< 0.001
Shopping ^t	0.17	[0.14-0.19]	< 0.001
Entertainment activities			
Other entertainment ^u	8.76	[2.03-37.7]	0.004
Visiting a zoo	4.17	[0.95-18.36]	0.059
Amusement park	2.48	[0.26-23.44]	0.428
Pub or Bar ^v	2.11	[0.83-5.39]	0.117
Park	0.71	[0.11-4.53]	0.72
Social club	0.6	[0.07-5.41]	0.647

*Adjusted for age, sex, ethnicity, region and index of multiple deprivation

- a. Working in warehouse settings – warehouse, haulage, distribution etc
- b. Working in close contact services – barbers, hairdressers, nail salons, tattoo studios and tanning salons, and any other services which require close contact
- c. Working in manufacturing and construction – textile, electronics, cars
- d. Working in hospitality – working in restaurants, food and drink outlets, lodging etc
- e. Education – working or attending
- f. Working in public transport- including buses, logistics, underground and trains
- g. Working in military – including the Navy, Army and Air Force

- h. Working in arts, or recreation – music, theatre, gyms, cinema, leisure centres
- i. Work related travel – including attending conferences, door-to-door sales, visiting clients
- j. Sports events - including football, cricket, rugby, tennis etc. matches, horse races and other sports events
- k. Personal exercise – e.g. at the gym, swimming, running clubs, casual exercise with friends
- l. Eating out – in different restaurants, cafes, pubs
- m. Charity and community – fundraising, volunteering, boot sale, corporate events, educational classes
- n. Private events – including weddings, funerals, parties, other social gatherings
- o. Mass gatherings – raves, gigs, festivals, protests and live music events
- p. Personal care - Visiting nail salons, hairdressers, barbers, tanning and tattoo studios
- q. Visiting family or friends – indoors or outdoors
- r. Travel - any international and domestic travel including trains and flights
- s. Health care - visiting healthcare services for non-COVID-19 reasons
- t. Shopping at supermarkets, local shops and pharmacies
- u. Other entertainment activities (non-specific)
- v. Going to pub or bar (entertainment, not e.g. eating out)

The model was fully adjusted for demographic variables. After adjusting for confounding, there remained strong statistical evidence of working in warehouse setting and being a COVID-19 case (OR 14.57, 95% C.I. 4.39-48.27, $P < 0.001$). There was also strong evidence of an association of being a COVID-19 case and working in healthcare, education, hospitality, social care, and construction being associated with increased odds of being a COVID-19 case. The analysis also showed strong evidence that cases were more likely to engage in 'Other Entertainment' activities than controls (OR 8.76, 95% C.I. 2.03-37.7, $P < 0.001$).

See Appendix 1 for regional MVA analyses.

3.5. Analysis of data related to the nature of an exposure setting (closed, crowded)

3.5.1 Analysis of closed settings

The multivariable analysis on closed settings was adjusted for age, sex, PHE region and IMD. There was weak evidence that working in a closed warehouse setting was associated with higher odds of being a COVID-19 case. However, across all other settings, there was no evidence to suggest that working or attending an activity in a closed space increased the odds of becoming a case.

Table 8: Multivariable analysis of the workplaces and activities in closed/not closed settings

Setting/Activity	Closed space	Odds Ratio	95% CI	P-value
Work and education				
Warehouse ^a	No	5.98	[1.58-22.69]	0.009
	Yes	50.3	[8.02-315.43]	< 0.001
Military ^g	No	3.18	[0.17-59.76]	0.439
	Yes	6.72	[0.88-51.37]	0.066
Hospitality ^d	No	3.74	[1.58-8.85]	0.003
	Yes	2.32	[1.29-4.17]	0.005
Arts and recreation ^h	No	3.72	[0.61-22.58]	0.154
	Yes	1.23	[0.55-2.76]	0.619
Healthcare	No	3.7	[2.13-6.43]	< 0.001
	Yes	2.57	[1.67-3.97]	< 0.001
Social care	No	3.19	[1.13-9.03]	0.029
	Yes	2.25	[1.06-4.79]	0.035
Construction ^c	No	3	[1.56-5.75]	0.001
	Yes	1.95	[1.19-3.19]	0.008
Close contact ^b	No	2.68	[0.83-8.65]	0.1
	Yes	2.14	[0.04-113.63]	0.708
Emergency services	No	0.68	[0.13-3.67]	0.652
	Yes	2.54	[0.95-6.78]	0.063
Education ^e	No	1.59	[1.08-2.33]	0.019
	Yes	2.23	[1.64-3.03]	< 0.001
Food industry	No	1.9	[0.58-6.3]	0.292
	Yes	1.36	[0.39-4.78]	0.633
Transport ^f	No	1.77	[0.74-4.2]	0.198
	Yes	1.59	[0.84-3.02]	0.157
Retail	No	1.3	[0.66-2.53]	0.448
	Yes	1.35	[0.82-2.23]	0.238
Work related travel ⁱ	No	0.92	[0.34-2.54]	0.879
	Yes	0.27	[0.09-0.84]	0.023
Leisure activities				
Entertainment	No	2.55	[1.15-5.62]	0.021
	Yes	2.72	[1.31-5.65]	0.007
Eating out ^l	No	1.23	[0.88-1.71]	0.234
	Yes	0.76	[0.59-0.98]	0.033
Private events ⁿ	No	1.2	[0.22-6.5]	0.83
	Yes	0.54	[0.13-2.3]	0.402
Personal exercise ^k	No	0.98	[0.67-1.44]	0.924
	Yes	1.04	[0.7-1.54]	0.84
Personal care ^p	No	0.89	[0.34-2.3]	0.805
	Yes	0.32	[0.18-0.57]	< 0.001
Travel ^r	No	0.84	[0.32-2.23]	0.731

Setting/Activity	Closed space	Odds Ratio	95% CI	P-value
Sports events ^l	Yes	0.33	[0.17-0.64]	0.001
	No	0.79	[0.35-1.8]	0.575
Visiting family ^q	Yes	0.35	[0.04-3.43]	0.367
	No	0.47	[0.3-0.73]	0.001
Community and charity events ^m	Yes	0.39	[0.28-0.53]	< 0.001
	No	0.45	[0.08-2.46]	0.359
Shopping ^t	Yes	0.41	[0.11-1.47]	0.169
	No	0.45	[0.37-0.56]	< 0.001
Visiting healthcare ^s	Yes	0.08	[0.07-0.1]	< 0.001
	No	0.43	[0.22-0.82]	0.011
	Yes	0.23	[0.14-0.36]	< 0.001

3.5.2 Analysis of crowded settings

The multivariable analysis on crowded settings was adjusted for age, sex, PHE region and IMD. There was no evidence that working or attending an activity in a crowded place increased the odds of becoming a case in nearly any of the settings. However, there was weak evidence that crowdedness in social care settings was associated with higher odds of being a COVID-19 case.

Table 9: Multivariable analysis of the workplaces and activities in crowded/non-crowded settings

Setting/Activity	Status	Odds Ratio	95% CI	P-value
Work and education				
Warehouse ^a	Not crowded	16.86	[4.45-63.83]	< 0.001
	Crowded	8.44	[0.77-93.06]	0.081
Military ^g	Not crowded	11.26	[1.15-110.44]	0.038
	Crowded	0.93	[0.04-20.41]	0.964
Social care	Not crowded	1.8	[0.97-3.35]	0.062
	Crowded	11.16	[1.8-69.2]	0.01
Emergency services	Not crowded	2.1	[0.84-5.27]	0.114
	Crowded	3.4	[0.43-27.09]	0.248
Healthcare	Not crowded	3.28	[2.24-4.79]	< 0.001
	Crowded	3.22	[1.51-6.88]	0.002
Hospitality ^d	Not crowded	2.31	[1.4-3.83]	0.001
	Crowded	2.89	[1.01-8.28]	0.049
Education ^e	Not crowded	2.85	[2.1-3.86]	< 0.001
	Crowded	0.99	[0.67-1.47]	0.974

Setting/Activity	Status	Odds Ratio	95% CI	P-value
Transport ^f	Not crowded	2.72	[1.46-5.05]	0.002
	Crowded	0.42	[0.11-1.55]	0.191
Construction ^c	Not crowded	2.49	[1.66-3.73]	< 0.001
	Crowded	0.82	[0.17-3.87]	0.798
Close contact ^b	Not crowded	2.05	[0.75-5.58]	0.159
	Crowded	1.62	[0.05-57.11]	0.791
Arts and recreation ^h	Not crowded	1.69	[0.76-3.8]	0.201
	Crowded	0.41	[0.08-2.25]	0.305
Retail	Not crowded	1.66	[1.03-2.68]	0.036
	Crowded	0.75	[0.35-1.6]	0.454
Food industry	Not crowded	1.46	[0.59-3.59]	0.41
	Crowded	0.56	[0.06-5.27]	0.615
Work related travel ⁱ	Not crowded	0.61	[0.29-1.29]	0.195
	Crowded	0.05	[0-1.72]	0.098
Leisure activities				
Entertainment	Not crowded	2.62	[1.47-4.68]	0.001
	Crowded	1.11	[0.25-4.95]	0.892
Private events ⁿ	Not crowded	0.82	[0.22-3.05]	0.766
	Crowded	1.18	[0.16-8.96]	0.873
Visiting family ^q	Not crowded	0.37	[0.28-0.49]	< 0.001
	Crowded	1.07	[0.45-2.51]	0.885
Sports events ^j	Not crowded	1.05	[0.47-2.34]	0.91
	Crowded	0.44	[0.04-4.93]	0.506
Eating out ^l	Not crowded	0.96	[0.78-1.2]	0.745
	Crowded	0.47	[0.27-0.82]	0.008
Personal care ^p	Not crowded	0.39	[0.23-0.67]	0.001
	Crowded	0.93	[0.22-3.89]	0.916
Personal exercise ^k	Not crowded	0.84	[0.64-1.12]	0.239
	Crowded	0.6	[0.27-1.36]	0.223
Community and charity events ^m	Not crowded	0.53	[0.17-1.64]	0.272
	Crowded	0.5	[0.01-27.27]	0.732
Travel ^r	Not crowded	0.47	[0.25-0.87]	0.017
	Crowded	0.38	[0.14-1.05]	0.063
Public events, mass gatherings ^o	Not crowded	0.29	[0.13-0.65]	0.003
	Crowded	0.06	[0-1.63]	0.095
Visiting healthcare ^s	Not crowded	0.29	[0.2-0.42]	< 0.001
	Crowded	0.14	[0.04-0.58]	0.006
Shopping ^t	Not crowded	0.2	[0.17-0.23]	< 0.001
	Crowded	0.1	[0.07-0.12]	< 0.001

*Adjusted for age, sex, region and index of multiple deprivation

a. Working in warehouse settings –warehouse, haulage, distribution etc

- b. Working in close contact services – barbers, hairdressers, nail salons, tattoo studios and tanning salons, and any other services which require close contact
- c. Working in manufacturing and construction – textile, electronics, cars
- d. Working in hospitality – working in restaurants, food and drink outlets, lodging etc
- e. Education – working or attending
- f. Working in public transport- including buses, logistics, underground and trains
- g. Working in military – including the Navy, Army and Air Force
- h. Working in arts, or recreation – music, theatre, gyms, cinema, leisure centres
- i. Work related travel – including attending conferences, door-to-door sales, visiting clients
- j. Sports events - including football, cricket, rugby, tennis etc. matches, horse races and other sports events
- k. Personal exercise – e.g. at the gym, swimming, running clubs, casual exercise with friends
- l. Eating out – in different restaurants, cafes, pubs
- m. Charity and community – fundraising, volunteering, bootsale, corporate events, educational classes
- n. Private events – including weddings, funerals, parties, other social gatherings
- o. Mass gatherings – raves, gigs, festivals, protests and live music events
- p. Personal care - Visiting nail salons, hairdressers, barbers, tanning and tattoo studios
- q. Visiting family or friends – indoors or outdoors
- r. Travel - any international and domestic travel including trains and flights
- s. Health care - visiting healthcare services for non-COVID-19 reasons
- t. Shopping at supermarkets, local shops and pharmacies
- u. Other entertainment activities (non-specific)
- v. Going to pub or bar (entertainment, not e.g. eating out)

3.6. Controlling for misclassification of outcome (COVID-19 case/control status)

All exposures (settings) that showed strong evidence of being positively associated with being a COVID-19 case in the main MVA continue to show evidence in any of the sensitivity/specificity scenarios tested (sensitivity as low as 90%, specificity as low as 90%, respectively and combined). These include working in warehouses and construction, healthcare, hospitality, and education, among others, as well as engaging in “other” entertainment activities. Please see Tables 10-12 below for full results.

Table 10: MVA results obtained employing an EM algorithm and an assumed sensitivity of 0.9 and specificity of 1.0

Setting/Activity	Odds Ratio	95% CI	P-value
Work and education			
Warehouse ^a	54.12	[3.27-894.74]	0.005
Military ^g	6.07	[0.65-56.95]	0.114
Healthcare ^s	3.35	[2.17-5.17]	< 0.001
Construction ^c	3.21	[1.88-5.47]	< 0.001
Social care	2.6	[1.22-5.52]	0.013
Close contact ^b	2.54	[0.6-10.66]	0.203
Emergency services	2.21	[0.74-6.66]	0.158
Education ^e	2.16	[1.59-2.93]	< 0.001
Hospitality ^d	2.05	[1.18-3.58]	0.011
Arts and recreation ^h	2.03	[0.81-5.08]	0.13
Transport ^f	1.74	[0.91-3.32]	0.096
Retail	1.17	[0.73-1.87]	0.521
Food industry	0.87	[0.28-2.66]	0.801
Work related travel ⁱ	0.39	[0.17-0.9]	0.027
Immigration	0	[0-0]	0.999

Setting/Activity	Odds Ratio	95% CI	P-value
Leisure activities			
Eating out ^l	0.86	[0.67-1.09]	0.21
Personal exercise ^k	0.81	[0.59-1.1]	0.175
Sports events ^j	0.72	[0.29-1.82]	0.49
Private events ⁿ	0.62	[0.18-2.18]	0.458
Personal care ^p	0.38	[0.22-0.66]	0.001
Community and charity events ^m	0.36	[0.1-1.33]	0.125
Travel ^r	0.36	[0.19-0.67]	0.001
Visiting family ^q	0.34	[0.25-0.46]	< 0.001
Visiting healthcare	0.22	[0.14-0.33]	< 0.001
Public events, mass gatherings ^o	0.14	[0.06-0.35]	< 0.001
Shopping ^t	0.12	[0.1-0.15]	< 0.001
Entertainment activities			
Other entertainment ^u	13.98	[1.55-125.9]	0.019
Visiting a zoo	6.47	[0.72-57.81]	0.095
Amusement park	2.32	[0.16-32.76]	0.533
Pub or Bar ^v	2.22	[0.71-6.93]	0.17
Park	0.52	[0.06-4.51]	0.552
Social club	0.41	[0.03-5.67]	0.505

*Adjusted for age, sex, region and index of multiple deprivation

- a. Working in warehouse settings – warehouse, haulage, distribution etc
- b. Working in close contact services – barbers, hairdressers, nail salons, tattoo studios and tanning salons, and any other services which require close contact
- c. Working in manufacturing and construction – textile, electronics, cars
- d. Working in hospitality – working in restaurants, food and drink outlets, lodging etc
- e. Education – working or attending
- f. Working in public transport- including buses, logistics, underground and trains
- g. Working in military – including the Navy, Army and Air Force
- h. Working in arts, or recreation – music, theatre, gyms, cinema, leisure centres
- i. Work related travel – including attending conferences, door-to-door sales, visiting clients
- j. Sports events - including football, cricket, rugby, tennis etc. matches, horse races and other sports events
- k. Personal exercise – e.g. at the gym, swimming, running clubs, casual exercise with friends
- l. Eating out – in different restaurants, cafes, pubs
- m. Charity and community – fundraising, volunteering, boot sale, corporate events, educational classes
- n. Private events – including weddings, funerals, parties, other social gatherings
- o. Mass gatherings – raves, gigs, festivals, protests and live music events
- p. Personal care - Visiting nail salons, hairdressers, barbers, tanning and tattoo studios
- q. Visiting family or friends – indoors or outdoors
- r. Travel - any international and domestic travel including trains and flights
- s. Health care - visiting healthcare services for non-COVID-19 reasons
- t. Shopping at supermarkets, local shops and pharmacies
- u. Other entertainment activities (non-specific)
- v. Going to pub or bar (entertainment, not e.g. eating out)

Table 11: MVA results obtained employing an EM algorithm and an assumed sensitivity of 1.0 and specificity of 0.9

Setting/Activity	Odds Ratio	95% CI	P-value
Work and education			
Warehouse ^a	24.09	[6.15-94.32]	< 0.001
Military ^g	9.38	[1.19-73.68]	0.033

Setting/Activity	Odds Ratio	95% CI	P-value
Healthcare ^s	3.79	[2.58-5.57]	< 0.001
Social care	3.02	[1.57-5.82]	0.001
Construction ^c	3.02	[1.93-4.74]	< 0.001
Hospitality ^d	2.72	[1.61-4.57]	< 0.001
Emergency services	2.63	[1.03-6.71]	0.043
Education ^e	2.3	[1.72-3.06]	< 0.001
Close contact ^b	2.22	[0.69-7.09]	0.18
Transport ^f	2.12	[1.17-3.81]	0.013
Arts and recreation ^h	1.97	[0.86-4.52]	0.109
Retail	1.43	[0.89-2.29]	0.141
Food industry	0.92	[0.3-2.78]	0.88
Work related travel ⁱ	0.38	[0.14-1]	0.05
Leisure activities	-		
Personal exercise ^k	0.93	[0.67-1.3]	0.681
Eating out ^l	0.86	[0.66-1.1]	0.233
Sports events ^j	0.72	[0.3-1.75]	0.47
Private events ⁿ	0.69	[0.18-2.67]	0.593
Travel ^r	0.41	[0.21-0.81]	0.01
Community and charity events ^m	0.38	[0.08-1.66]	0.197
Personal care ^p	0.35	[0.18-0.71]	0.003
Visiting family ^q	0.34	[0.24-0.48]	< 0.001
Visiting healthcare	0.18	[0.1-0.32]	< 0.001
Public events, mass gatherings ^o	0.14	[0.05-0.4]	< 0.001
Shopping ^t	0.12	[0.1-0.14]	< 0.001
Entertainment activities	-		
Other entertainment ^u	14.55	[2.68-79.05]	0.002
Visiting a zoo	5.86	[1.04-33.01]	0.045
Amusement park	3.63	[0.27-49.51]	0.334
Pub or Bar ^v	2.56	[0.84-7.78]	0.096
Park	0.71	[0.08-6.64]	0.761
Social club	0.55	[0.03-9.32]	0.677

*Adjusted for age, sex, region and index of multiple deprivation

- a. Working in warehouse settings – warehouse, haulage, distribution etc
- b. Working in close contact services – barbers, hairdressers, nail salons, tattoo studios and tanning salons, and any other services which require close contact
- c. Working in manufacturing and construction – textile, electronics, cars
- d. Working in hospitality – working in restaurants, food and drink outlets, lodging etc
- e. Education – working or attending
- f. Working in public transport- including buses, logistics, underground and trains
- g. Working in military – including the Navy, Army and Air Force
- h. Working in arts, or recreation – music, theatre, gyms, cinema, leisure centres
- i. Work related travel – including attending conferences, door-to-door sales, visiting clients
- j. Sports events - including football, cricket, rugby, tennis etc. matches, horse races and other sports events
- k. Personal exercise – e.g. at the gym, swimming, running clubs, casual exercise with friends
- l. Eating out – in different restaurants, cafes, pubs
- m. Charity and community – fundraising, volunteering, boot sale, corporate events, educational classes
- n. Private events – including weddings, funerals, parties, other social gatherings
- o. Mass gatherings – raves, gigs, festivals, protests and live music events
- p. Personal care - Visiting nail salons, hairdressers, barbers, tanning and tattoo studios
- q. Visiting family or friends – indoors or outdoors
- r. Travel - any international and domestic travel including trains and flights
- s. Health care - visiting healthcare services for non-COVID-19 reasons
- t. Shopping at supermarkets, local shops and pharmacies

- u. *Other entertainment activities (non-specific)*
- v. *Going to pub or bar (entertainment, not e.g. eating out)*

Table 12: MVA results obtained employing an EM algorithm and an assumed sensitivity of 0.9 and specificity of 0.9

Setting/Activity	Odds Ratio	95% CI	P-value
Work and education			
Warehouse ^a	95.86	[6.27-1466.02]	0.001
Military ^g	8.92	[0.75-106.52]	0.084
Healthcare ^s	4.58	[2.74-7.65]	< 0.001
Construction ^c	4.15	[2.2-7.81]	< 0.001
Social care	3.52	[1.5-8.23]	0.004
Close contact ^b	3.16	[0.62-15.98]	0.164
Emergency services	2.78	[0.78-9.88]	0.113
Education ^e	2.67	[1.85-3.87]	< 0.001
Hospitality ^d	2.6	[1.36-4.94]	0.004
Arts and recreation ^h	2.53	[0.88-7.31]	0.087
Transport ^f	2.17	[1.03-4.57]	0.042
Retail	1.38	[0.78-2.45]	0.264
Food industry	0.81	[0.2-3.2]	0.759
Work related travel ⁱ	0.32	[0.11-0.92]	0.035
Immigration	0	[0-0]	0.999
Leisure activities			
Personal exercise ^k	0.88	[0.6-1.31]	0.537
Eating out ^l	0.82	[0.6-1.11]	0.197
Sports events ^l	0.68	[0.22-2.12]	0.504
Private events ⁿ	0.62	[0.13-2.86]	0.54
Community and charity events ^m	0.41	[0.08-2.19]	0.297
Travel ^f	0.34	[0.15-0.75]	0.007
Personal care ^p	0.33	[0.15-0.7]	0.004
Visiting family ^q	0.27	[0.18-0.41]	< 0.001
Visiting healthcare	0.14	[0.07-0.26]	< 0.001
Public events, mass gatherings ^o	0.09	[0.03-0.31]	< 0.001
Shopping ^t	0.08	[0.06-0.11]	< 0.001
Entertainment activities			
Other entertainment ^u	20.57	[1.97-215.21]	0.012
Visiting a zoo	10.37	[0.88-121.63]	0.063
Amusement park	3.12	[0.16-60.15]	0.451
Pub or Bar ^v	2.97	[0.74-11.88]	0.123
Park	0.62	[0.05-7.37]	0.709
Social club	0.48	[0.02-9.79]	0.633

*Adjusted for age, sex, region and index of multiple deprivation

- Working in warehouse settings – warehouse, haulage, distribution etc
- Working in close contact services – barbers, hairdressers, nail salons, tattoo studios and tanning salons, and any other services which require close contact
- Working in manufacturing and construction – textile, electronics, cars
- Working in hospitality – working in restaurants, food and drink outlets, lodging etc
- Education – working or attending
- Working in public transport- including buses, logistics, underground and trains
- Working in military – including the Navy, Army and Air Force
- Working in arts, or recreation – music, theatre, gyms, cinema, leisure centres
- Work related travel – including attending conferences, door-to-door sales, visiting clients

- j. Sports events - including football, cricket, rugby, tennis etc. matches, horse races and other sports events
- k. Personal exercise – e.g. at the gym, swimming, running clubs, casual exercise with friends
- l. Eating out – in different restaurants, cafes, pubs
- m. Charity and community – fundraising, volunteering, bootsale, corporate events, educational classes
- n. Private events – including weddings, funerals, parties, other social gatherings
- o. Mass gatherings – raves, gigs, festivals, protests and live music events
- p. Personal care - Visiting nail salons, hairdressers, barbers, tanning and tattoo studios
- q. Visiting family or friends – indoors or outdoors
- r. Travel - any international and domestic travel including trains and flights
- s. Health care - visiting healthcare services for non-COVID-19 reasons
- t. Shopping at supermarkets, local shops and pharmacies
- u. Other entertainment activities (non-specific)
- v. Going to pub or bar (entertainment, not e.g. eating out)

3.7. Controlling for mode of questionnaire completion (online/telephone)

Of the 2,000 cases, 879 (44%) were interviewed on the phone. 1,121 (56%) cases completed the enhanced contact tracing questionnaire online (self-completion). The following analysis explores whether the cases who self-completed the enhanced contact tracing questionnaire differed from the cases who were interviewed. Tables 13 and 14 show the results from multinomial regression analyses. The model is adjusted for demographic variables age, sex, ethnicity as well as deprivation level.

Mode of completion did not have a broad effect on the association between work and education settings and the outcome of being a COVID-19 case. One setting that was affected was working in emergency services, which had a substantially higher relative risk ratios (RRR) for cases interviewed on the phone compared to self-completion online.

Mode of completion did not affect associations between leisure activities and the outcome of being a COVID-19 case.

Table 13: Multinomial analysis of work and education settings, by type of completion

Setting/Activity	Relative Risk Ratio	95% CI	P-value
Self-completed			
Warehouse ^a	17.53	[5.16-59.59]	< 0.001
Military ^g	5.18	[0.86-31.3]	0.073
Healthcare ^s	3.81	[2.67-5.42]	< 0.001
Social care	3.6	[1.92-6.77]	< 0.001
Construction ^c	3.02	[1.99-4.6]	< 0.001
Hospitality ^d	2.48	[1.52-4.04]	< 0.001
Close contact ^b	2.42	[0.88-6.68]	0.088
Transport ^f	2.02	[1.15-3.55]	0.014
Education ^e	1.88	[1.41-2.51]	< 0.001
Arts and recreation ^h	1.72	[0.74-3.99]	0.207
Emergency services	1.46	[0.44-4.83]	0.533
Retail	1.24	[0.77-1.98]	0.37
Food industry	1.17	[0.44-3.11]	0.758
Work related travel ⁱ	0.11	[0.03-0.48]	0.003
Immigration	0	[0-0]	0.991
Telephone interview			

Setting/Activity	Relative Risk Ratio	95% CI	P-value
Warehouse ^a	22.09	[6.58-74.15]	< 0.001
Military ^g	8.14	[1.67-39.73]	0.01
Emergency services	4.77	[2.11-10.79]	< 0.001
Construction ^c	3.95	[2.69-5.82]	< 0.001
Healthcare ^s	3.39	[2.45-4.69]	< 0.001
Social care	2.83	[1.54-5.19]	0.001
Hospitality ^d	2.55	[1.63-3.99]	< 0.001
Arts and recreation ^h	2.44	[1.23-4.86]	0.011
Transport ^f	2.42	[1.43-4.09]	0.001
Close contact ^b	2.36	[0.92-6.08]	0.075
Education ^e	2.33	[1.84-2.96]	< 0.001
Retail	1.35	[0.89-2.03]	0.155
Food industry	0.73	[0.25-2.15]	0.573
Work related travel ⁱ	0.39	[0.18-0.85]	0.018
Immigration	0	[0-0]	0.989

*Adjusted for age, sex, and index of multiple deprivation

- a. Working in warehouse settings – warehouse, haulage, distribution etc
- b. Working in close contact services – barbers, hairdressers, nail salons, tattoo studios and tanning salons, and any other services which require close contact
- c. Working in manufacturing and construction – textile, electronics, cars
- d. Working in hospitality – working in restaurants, food and drink outlets, lodging etc
- e. Education – working or attending
- f. Working in public transport- including buses, logistics, underground and trains
- g. Working in military – including the Navy, Army and Air Force
- h. Working in arts, or recreation – music, theatre, gyms, cinema, leisure centres
- i. Work related travel – including attending conferences, door-to-door sales, visiting clients
- j. Sports events - including football, cricket, rugby, tennis etc. matches, horse races and other sports events
- k. Personal exercise – e.g. at the gym, swimming, running clubs, casual exercise with friends
- l. Eating out – in different restaurants, cafes, pubs
- m. Charity and community – fundraising, volunteering, bootsale, corporate events, educational classes
- n. Private events – including weddings, funerals, parties, other social gatherings
- o. Mass gatherings – raves, gigs, festivals, protests and live music events
- p. Personal care - Visiting nail salons, hairdressers, barbers, tanning and tattoo studios
- q. Visiting family or friends – indoors or outdoors
- r. Travel - any international and domestic travel including trains and flights
- s. Health care - visiting healthcare services for non-COVID-19 reasons
- t. Shopping at supermarkets, local shops and pharmacies
- u. Other entertainment activities (non-specific)
- v. Going to pub or bar (entertainment, not e.g. eating out)

Table 14: Multinomial analysis of leisure activity settings, by type of completion

Setting/Activity	Relative Risk Ratio	95% CI	P-value
Self-completed			
Entertainment	2.49	[1.33-4.65]	0.004
Private events ⁿ	1.15	[0.32-4.17]	0.828
Eating out ^l	0.79	[0.6-1.02]	0.074
Personal exercise ^k	0.63	[0.44-0.91]	0.014
Sports events ^j	0.62	[0.22-1.76]	0.368
Travel ^r	0.47	[0.23-0.96]	0.038
Visiting family ^q	0.33	[0.23-0.47]	< 0.001
Visiting healthcare	0.31	[0.19-0.5]	< 0.001
Personal care ^p	0.28	[0.12-0.62]	0.002
Shopping ^t	0.2	[0.16-0.24]	< 0.001
Community and charity events ^m	0.19	[0.02-1.51]	0.116
Public events, mass gatherings ^o	0.04	[0.01-0.3]	0.002
Telephone interview			
Entertainment	2.33	[1.33-4.09]	0.003
Personal exercise ^k	0.99	[0.74-1.34]	0.966
Sports events ^j	0.96	[0.44-2.11]	0.92
Eating out ^l	0.93	[0.74-1.17]	0.518
Community and charity events ^m	0.6	[0.18-2.01]	0.406
Personal care ^p	0.48	[0.28-0.85]	0.011
Travel ^r	0.45	[0.24-0.86]	0.016
Visiting family ^q	0.41	[0.3-0.55]	< 0.001
Private events ⁿ	0.37	[0.07-1.89]	0.23
Public events, mass gatherings ^o	0.33	[0.14-0.76]	0.01
Visiting healthcare	0.17	[0.1-0.29]	< 0.001
Shopping ^t	0.12	[0.1-0.15]	< 0.001

*Adjusted for age, sex, and index of multiple deprivation

- a. Working in warehouse settings – warehouse, haulage, distribution etc
- b. Working in close contact services – barbers, hairdressers, nail salons, tattoo studios and tanning salons, and any other services which require close contact
- c. Working in manufacturing and construction – textile, electronics, cars
- d. Working in hospitality – working in restaurants, food and drink outlets, lodging etc
- e. Education – working or attending
- f. Working in public transport- including buses, logistics, underground and trains
- g. Working in military – including the Navy, Army and Air Force
- h. Working in arts, or recreation – music, theatre, gyms, cinema, leisure centres
- i. Work related travel – including attending conferences, door-to-door sales, visiting clients
- j. Sports events - including football, cricket, rugby, tennis etc. matches, horse races and other sports events
- k. Personal exercise – e.g. at the gym, swimming, running clubs, casual exercise with friends
- l. Eating out – in different restaurants, cafes, pubs
- m. Charity and community – fundraising, volunteering, bootsale, corporate events, educational classes
- n. Private events – including weddings, funerals, parties, other social gatherings
- o. Mass gatherings – raves, gigs, festivals, protests and live music events
- p. Personal care - Visiting nail salons, hairdressers, barbers, tanning and tattoo studios
- q. Visiting family or friends – indoors or outdoors
- r. Travel - any international and domestic travel including trains and flights
- s. Health care - visiting healthcare services for non-COVID-19 reasons
- t. Shopping at supermarkets, local shops and pharmacies
- u. Other entertainment activities (non-specific)
- v. Going to pub or bar (entertainment, not e.g. eating out)

4 Discussion

The study showed that there was strong statistical evidence that working in warehouse settings and construction, education and hospitality, as well as health and social care were associated with increased odds of being COVID-19 case. Additionally, there was statistical evidence that engaging in “other” entertainment activities like going to theme parks and golf clubs was associated increased odds of being a case. There was no evidence that the nature of the activity setting (closed space or crowdedness) increased the odds of being a COVID-19 case. These results were aligned with our previous case-control studies conducted in early and late September. When adjusting for misclassification arising from imperfect sensitivity and specificity of RT-PCR test method, the settings that showed evidence in the main MVA remained associated with increased odds of being a COVID-19 case.

The results presented in this report are mostly consistent with the previous iterations of this study (see Appendix 4), which showed evidence of working in healthcare, social care, warehouse, and education being associated with increased odds of disease. The previous studies also provided similar evidence on leisure activities.

Overall, most of the findings in this study are also consistent with existing literature. The findings regarding hospitality are in line with a study conducted by the CDC (Fisher et al., 2020; ECDC, 2020) that showed substantially increased risks associated with visiting a restaurant, bar, or coffee shop, where masks cannot be effectively worn when consuming food or drink in hospitality settings. This might explain the increased risk of COVID-19 infection for individuals working in hospitality. Visiting a pub or bar was not found to be associated with being a COVID-19 case in this iteration of the study, but was associated in the second study (see Appendix 4). This might possibly be due to the fact that the majority of the cases were sampled from areas where, at the time of the interviews, Tier 3 restrictions were in place. This could have been associated with the signal for pubs and bars being weaker during the third study.

The findings on associations between higher odds of illness and working in manufacturing, construction, and warehouse settings highlight the increased risk associated with these settings in the context of COVID-19. They are also consistent with the previous iterations of the study, which all found strong evidence for an association between working in warehouse setting and being a COVID-19 case. These results are also aligned with the current evidence on most commonly reported outbreak settings (ECDC, 2020), which shows that COVID-19 clusters being reported in occupational sectors deemed as critical or essential occupational sectors. It is plausible that in these settings adhering to social distancing may be challenging, and workers may also share living spaces or rides to work. To investigate whether this signal might be caused by increased access to testing provided by specific

warehouse companies, we examined the type of companies cases reported to work for in warehouse settings. These included a very diverse sets of employments from self-employed to small or mid-sized companies to large companies, suggesting that the association between warehouse settings and being a COVID-19 case might not be driven by increased access to testing provided by specific warehouse companies.

Our findings on working in the educational sector/attending education institutions being associated with increased odds of being a COVID-19 case are consistent with the accumulating evidence of the risk of SARS-CoV-2 transmission in educational settings (Ismail et al., 2020). The previous studies did not provide strong evidence of a positive association between education settings and being a COVID-19 case. This may be due to the academic year now having started and students and staff possibly attending in person, which was not the case when the previous iterations of this study were run. This trend is also supported by the fact that universities were one of the most commonly reported exposure settings among cases who reported working in educational settings in the third study.

The findings on cases having higher odds of working in healthcare and social care are similar to previous iterations of the study. They are also aligned with increasing literature on the risk of transmission of SARS-CoV-2 in healthcare settings (Nguyen et al., 2020; Leclerc et al., 2020; ECDC, 2020). However, in this instance increased access to testing might be one of the drivers of the association obtained.

While there was very little of evidence the nature of the setting where the activity took place (crowdedness or closed space) was associated with increased odds of becoming a case. The number of those exposed in both cases and controls were relatively low, which was also reflected in the wide confidence intervals for many of the estimates. It is plausible that increased crowdedness and closed spaces would increase the risk of SARS-CoV-2 transmission, however, the data presented in this study does not allow us to provide evidence on this.

4.2 Limitations

Limitations of the present study can be grouped into categories relating to means of data collection, sampling and selection bias, misclassification bias and confounding.

4.2.1. Means of data collection

While cases completed the enhanced contact tracing questionnaire either online (self-completion) or on the telephone, data on all controls was collected via an online questionnaire. This might affect how comparable the two samples are. Mode of completion did not have a broad effect on associations between exposures (settings/activities) and the outcome of being a COVID-19 case except for working in emergency services.

There was some evidence of excessive responses among the controls. Around 7% of controls reported for example excessive number of work or education exposures in the past 7 days. These controls were excluded from analyses. This may be an indicator of poor quality responses in controls which may explain why some of the controls consistently reported higher number of exposures than cases. High number of reported exposures among controls can also be a reflection of difference in motivation to complete the questionnaire. Controls, who received payment once completed, may have been more motivated to report more activities than cases.

4.2.2. Sampling and selection bias

The study population only included adults and thus does not investigate the impact of the exposures commonly associated with children. Cases were randomly sampled from cases aged 18+ in the NHS T&T system which contains exposure information for all COVID-19 cases in England. The controls were sampled from a pool of volunteers using Market Research Panels, which most likely introduced selection bias for controls. People registered on Market Research Panels are likely to differ from the general population. Their age and sex distribution might differ and they might be more or less likely to engage in social activities that are of interest in this study. Furthermore, people registered on the Market Research Panel who chose to participate in the present study might differ from those registered on the Panel who did not engage with the study. It is therefore likely that our sample of controls does not accurately represent the adult population of England. Effect estimates (ORs) are likely to be biased, but the direction (reduction or increase) in which they are biased cannot easily be discerned due to a multitude of possibly competing factors.

4.2.3. Misclassification bias (case/control status)

We explored the possible effect of misclassification of case/control status, which might have occurred to a small degree. While this is unlikely to be a substantial limitation of the study, cases included in the present study might have received a false positive SARS-CoV-2 test result. Controls might be infected with COVID-19 but not show symptoms, not been tested, or received a false negative result.

The potential misclassification arising from lower sensitivity and specificity of RT-PCR testing method may result in participants with a false positive result be classified as cases and participants with an asymptomatic COVID-19 infection/false negative test result as controls. This would likely dilute the effect estimates (ORs) towards the null (OR 1.0), which is likely to lead to an underestimation of risk (OR) for any exposure. We explored this bias in sensitivity analyses accounting for misclassification of outcome status. The results obtained from these sensitivity analyses show higher effect measure estimates, which would suggest that the the

results obtained from the main multivariable analysis are more likely to be underestimations than overestimations.

4.2.4. Misclassification of exposure

Cases reported less exposures compared to controls. The median number of exposures reported by cases was 2 (range 1-20), whereas for controls this was 5 (range 1- 68). One possible explanation for this is social desirability bias (Maccoby, 1954), i.e., the motivation of study participants to show no characteristics or behaviours that might be deemed socially unacceptable. In reference to the current pandemic, behaviours such as going out to bars and restaurants and attending mass gatherings might be viewed as socially unacceptable by a substantial proportion of the population since they might be seen to contribute to further the spread of COVID-19. This might lead cases to (falsely) not report these exposures, leading to a misclassification of their exposure status. However, another possible driver might be that while we excluded household contacts of confirmed COVID-19 cases from our analysis, some cases might still have arisen as contacts of previously confirmed COVID-19 cases. Additionally, cases were asked to provide more details than controls if they reported exposure to a specific setting for public health action purposes. This additional level of detail required in the context of each activity might have also motivated cases to underreport the number of their activities compared to controls, leading to a differential misclassification of the exposure.

Regarding the exposures of controls, the fact that they were paid to fill out the online questionnaire might have motivated some to do it in great detail, but others to rush through it to obtain the financial compensation as quickly as possible. This may have lead to exposure information not being recorded accurately.

In addition to that recall bias might play a role. Cases might be more likely to recall exposures in trying to find the source of their infection than controls who might be less motivated to recall their exposures accurately, partly counteracting the above effect of social desirability bias.

It is difficult to account for the issues around misclassification of exposure for both cases and controls. Assuming that the underreporting of exposures (social desirability bias/contacts of previous cases) had a greater effect on the data than the failure/motivation of controls to recall their exposures (recall bias), as evidenced by the lower number of median exposures reported by cases, the effect estimates (ORs) are likely to have been biased towards the null (OR 1.0). This is likely to lead to an underestimation of risk (measured by odds ratio, OR) for any exposure.

4.2.5. Confounding

Multivariable analyses controlled for confounding by all other exposures in the models in addition to demographic variables. Still, residual confounding is likely to affect these observed associations between the exposures (settings) and being a COVID-19 case. For example, while demographics of cases and controls were broadly similar, IMD score was missing for 11% of controls and the location of residence was not perfectly matched between the groups, leading to residual confounding even after running a multivariate model.

In addition, present analyses could only control for confounding by exposures that data were available on – a sample that is (in varying degrees) affected by selection and misclassification bias and that might not contain data on any potential confounder conceivable.

4.2.6. Future improvements

We will aim to recruit an additional control group to examine the effects of selection bias among controls. Multivariable analyses controlling for household size of cases and controls are also planned.

5. Conclusion

The study shows evidence that certain occupational groups are associated with increased odds of being COVID-19 case.

In terms of non-work related factors, engaging in entertainment activities was associated with higher odds of being a COVID-19 case. These findings should be viewed in the context of the limitations of the study. However, the results presented in this study show consistency with the existing evidence from other studies and the previous iterations of this case-control study. This information should be used together with other insights on settings and activities associated with transmission of COVID-19 to inform control measures in England.

References

- Arevalo, I. (n.d.). FALSE-NEGATIVE RESULTS OF INITIAL RT-PCR ASSAYS FOR COVID-19: A SYSTEMATIC REVIEW, 26.
<https://www.medrxiv.org/content/medrxiv/early/2020/04/21/2020.04.16.20066787.full.pdf>
- Böger, B., Fachi, M. M., Vilhena, R. O., Cobre, A. F., Tonin, F. S., & Pontarolo, R. (2020). Systematic review with meta-analysis of the accuracy of diagnostic tests for COVID-19. *American Journal of Infection Control*. <https://doi.org/10.1016/j.ajic.2020.07.011>
- European Centers for Disease Control. (2020). COVID-19 clusters and outbreaks in occupational settings in the EU/EEA and the UK, 11 August 2020. European Centers for Disease Control. *TECHNICAL REPORT*, 17. <https://www.ecdc.europa.eu/sites/default/files/documents/COVID-19-in-occupational-settings.pdf>
- Epidemiological changes on the Isle of Wight after the launch of the NHS Test and Trace programme: a preliminary analysis - The Lancet Digital Health. (n.d.). Retrieved October 28, 2020, from [https://www.thelancet.com/journals/landig/article/PIIS2589-7500\(20\)30241-7/fulltext](https://www.thelancet.com/journals/landig/article/PIIS2589-7500(20)30241-7/fulltext)
- Firth, D. (1993). Bias Reduction of Maximum Likelihood Estimates. *Biometrika*, 80(1), 27–38.
<https://doi.org/10.2307/2336755>
- Fisher, K. A. (2020). Community and Close Contact Exposures Associated with COVID-19 Among Symptomatic Adults ≥18 Years in 11 Outpatient Health Care Facilities — United States, July 2020. *MMWR. Morbidity and Mortality Weekly Report*, 69. <https://doi.org/10.15585/mmwr.mm6936a5>
- G, H., & M, S. (2002, August 30). A solution to the problem of separation in logistic regression.
<https://doi.org/10.1002/sim.1047>
- He, J.-L., Luo, L., Luo, Z.-D., Lyu, J.-X., Ng, M.-Y., Shen, X.-P., & Wen, Z. (2020). Diagnostic performance between CT and initial real-time RT-PCR for clinically suspected 2019 coronavirus disease (COVID-19) patients outside Wuhan, China. *Respiratory Medicine*, 168, 105980.
<https://doi.org/10.1016/j.rmed.2020.105980>

- Ismail, S. A., Saliba, V., Bernal, J. A. L., Ramsay, M. E., & Ladhani, S. N. (2020). SARS-CoV-2 infection and transmission in educational settings: cross-sectional analysis of clusters and outbreaks in England. *MedRxiv*, 2020.08.21.20178574. <https://doi.org/10.1101/2020.08.21.20178574>
- Kendall, M., Milsom, L., Abeler-Dörner, L., Wymant, C., Ferretti, L., Briers, M., ... Fraser, C. (2020). Epidemiological changes on the Isle of Wight after the launch of the NHS Test and Trace programme: a preliminary analysis. *The Lancet Digital Health*. [https://doi.org/10.1016/S2589-7500\(20\)30241-7](https://doi.org/10.1016/S2589-7500(20)30241-7)
- Leclerc, Q. J., Fuller, N. M., Knight, L. E., CMMID COVID-19 Working Group, Funk, S., & Knight, G. M. (2020). What settings have been linked to SARS-CoV-2 transmission clusters? *Wellcome Open Research*, 5, 83. <https://doi.org/10.12688/wellcomeopenres.15889.2>
- Lei, H., Xu, X., Xiao, S., Wu, X., & Shu, Y. (2020). Household transmission of COVID-19—a systematic review and meta-analysis. *Journal of Infection*, 0(0). <https://doi.org/10.1016/j.jinf.2020.08.033>
- Luo, L., Liu, D., Liao, X., Wu, X., Jing, Q., Zheng, J., ... Mao, C. (2020). Contact Settings and Risk for Transmission in 3410 Close Contacts of Patients With COVID-19 in Guangzhou, China. *Annals of Internal Medicine*. <https://doi.org/10.7326/M20-2671>
- Maccoby, E. E. (1954). Why Do Children Watch Television? *The Public Opinion Quarterly*, 18(3), 239–244.
- Magder, L. S., & Hughes, J. P. (1997). Logistic Regression When the Outcome Is Measured with Uncertainty. *American Journal of Epidemiology*, 146(2), 195–203. <https://doi.org/10.1093/oxfordjournals.aje.a009251>
- Nguyen, L. H., Drew, D. A., Graham, M. S., Joshi, A. D., Guo, C.-G., Ma, W., ... Zhang, F. (2020a). Risk of COVID-19 among front-line health-care workers and the general community: a prospective cohort study. *The Lancet Public Health*, 5(9), e475–e483. [https://doi.org/10.1016/S2468-2667\(20\)30164-X](https://doi.org/10.1016/S2468-2667(20)30164-X)
- Office of National Statistics (ONS). (2020). Coronavirus (COVID-19) Infection Survey, UK: 23 October 2020. <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveypilot/23october2020>

- Sethuraman, N., Jeremiah, S. S., & Ryo, A. (2020). Interpreting Diagnostic Tests for SARS-CoV-2. *JAMA*, 323(22), 2249. <https://doi.org/10.1001/jama.2020.8259>
- Tahamtan, A., & Ardebili, A. (2020). Real-time RT-PCR in COVID-19 detection: issues affecting the results. *Expert Review of Molecular Diagnostics*, 20(5), 453–454. <https://doi.org/10.1080/14737159.2020.1757437>
- Watson, J., Whiting, P. F., & Brush, J. E. (2020). Interpreting a covid-19 test result, 8.
- World Health Organization. (2020). Laboratory testing for coronavirus disease 2019 (COVID-19) in suspected human cases: interim guidance, 2 March 2020. World Health Organization. <https://apps.who.int/iris/bitstream/handle/10665/331329/WHO-COVID-19-laboratory-2020.4-eng.pdf?sequence=1&isAllowed=y>
- Wölfel, R., Corman, V. M., Guggemos, W., Seilmaier, M., Zange, S., Müller, M. A., ... Wendtner, C. (2020). Virological assessment of hospitalized patients with COVID-2019. *Nature*, 581(7809), 465–469. <https://doi.org/10.1038/s41586-020-2196-x>
- Xie, C., Jiang, L., Huang, G., Pu, H., Gong, B., Lin, H., ... Yang, Z. (2020). Comparison of different samples for 2019 novel coronavirus detection by nucleic acid amplification tests. *International Journal of Infectious Diseases*, 93, 264–267. <https://doi.org/10.1016/j.ijid.2020.02.050>
- Yu, F., Yan, L., Wang, N., Yang, S., Wang, L., Tang, Y., ... Zhang, F. (2020). Quantitative Detection and Viral Load Analysis of SARS-CoV-2 in Infected Patients. *Clinical Infectious Diseases*, 71(15), 793–798. <https://doi.org/10.1093/cid/ciaa345>

Appendix 1.

Multivariable analyses by type of exposure and region.

Table A1-1: London

Setting/Activity	Odds Ratio	95% CI	P-value
Work and education			
Construction ^c	3.83	[0.72-20.35]	0.115
Arts and recreation ^h	3.62	[0.48-27.28]	0.211
Education ^e	2.44	[1-5.96]	0.051
Transport ^f	2.06	[0.47-9]	0.335
Healthcare	1.54	[0.48-4.94]	0.467
Hospitality ^d	1.26	[0.32-4.99]	0.743
Emergency services	0.91	[0.1-8.28]	0.934
Food industry	0.87	[0.02-39.83]	0.943
Close contact ^b	0.86	[0.01-51.27]	0.944
Work related travel ⁱ	0.64	[0.08-5.15]	0.674
Social care	0.55	[0.02-15]	0.722
Retail	0.43	[0.12-1.56]	0.2
Leisure activities			
Private events ⁿ	2.08	[0.08-54.48]	0.66
Personal care ^p	1.37	[0.35-5.38]	0.648
Eating out ^l	1.3	[0.69-2.47]	0.421
Sports events ^j	0.7	[0.11-4.67]	0.714
Personal exercise ^k	0.68	[0.29-1.6]	0.373
Visiting family ^q	0.5	[0.21-1.17]	0.11
Community and charity events ^m	0.47	[0.02-12.95]	0.652
Travel ^r	0.32	[0.06-1.69]	0.179
Shopping ^t	0.21	[0.13-0.35]	< 0.001
Visiting healthcare ^s	0.15	[0.03-0.66]	0.012
Public events, mass gatherings ^o	0.03	[0-0.55]	0.018
Entertainment activities			
Pub or Bar ^v	0.87	[0.12-6.28]	0.89
Park	0.68	[0.05-9.38]	0.773
Amusement park	0.58	[0-320.49]	0.867
Visiting a zoo	0.5	[0.02-13.28]	0.677

Note: Some regressors were automatically excluded due to multicollinearity.

- Working in warehouse settings – warehouse, haulage, distribution etc
- Working in close contact services – barbers, hairdressers, nail salons, tattoo studios and tanning salons, and any other services which require close contact
- Working in manufacturing and construction – textile, electronics, cars
- Working in hospitality – working in restaurants, food and drink outlets, lodging etc
- Education – working or attending
- Working in public transport- including buses, logistics, underground and trains
- Working in military – including the Navy, Army and Air Force

- h. Working in arts, or recreation – music, theatre, gyms, cinema, leisure centres
- i. Work related travel – including attending conferences, door-to-door sales, visiting clients
- j. Sports events - including football, cricket, rugby, tennis etc. matches, horse races and other sports events
- k. Personal exercise – e.g. at the gym, swimming, running clubs, casual exercise with friends
- l. Eating out – in different restaurants, cafes, pubs
- m. Charity and community – fundraising, volunteering, bootsale, corporate events, educational classes
- n. Private events – including weddings, funerals, parties, other social gatherings
- o. Mass gatherings – raves, gigs, festivals, protests and live music events
- p. Personal care - Visiting nail salons, hairdressers, barbers, tanning and tattoo studios
- q. Visiting family or friends – indoors or outdoors
- r. Travel - any international and domestic travel including trains and flights
- s. Health care - visiting healthcare services for non-COVID-19 reasons
- t. Shopping at supermarkets, local shops and pharmacies
- u. Other entertainment activities (non-specific)
- v. Going to pub or bar (entertainment, not e.g. eating out)

Table A1-2: Midlands (incl. East and West Midlands)

Setting/Activity	Odds Ratio	95% CI	P-value
Work and education			
Warehouse ^a	38.14	[1.98-735.4]	0.016
Social care	6.64	[1.3-33.95]	0.023
Healthcare	5.03	[1.75-14.42]	0.003
Emergency services	4.21	[0.79-22.35]	0.091
Arts and recreation ^h	4.11	[0.39-43.36]	0.24
Construction ^c	3.98	[1.41-11.19]	0.009
Food industry	3.91	[0.38-39.73]	0.249
Education ^e	2.56	[1.37-4.79]	0.003
Hospitality ^d	1.54	[0.4-5.97]	0.534
Transport ^f	1.5	[0.41-5.45]	0.537
Retail	1.21	[0.47-3.14]	0.693
Close contact ^b	0.7	[0.08-6.03]	0.741
Work related travel ⁱ	0.59	[0.09-3.85]	0.58
Leisure activities			
Private events ⁿ	1.48	[0.05-43.12]	0.821
Sports events ^j	0.92	[0.11-7.61]	0.942
Eating out ^l	0.79	[0.46-1.34]	0.38
Personal exercise ^k	0.68	[0.34-1.38]	0.288
Public events, mass gatherings ^o	0.65	[0.1-4.19]	0.654
Travel ^r	0.57	[0.15-2.17]	0.408
Visiting healthcare ^s	0.57	[0.23-1.39]	0.214
Visiting family ^q	0.39	[0.21-0.73]	0.003
Personal care ^p	0.32	[0.1-1.02]	0.054
Shopping ^t	0.16	[0.11-0.24]	< 0.001
Community and charity events ^m	0.06	[0-3.45]	0.173
Entertainment activities			
Other entertainment ^u	8.21	[0.28-239.38]	0.221
Visiting a zoo	2.31	[0.09-58.93]	0.614

Setting/Activity	Odds Ratio	95% CI	P-value
Social club	2.3	[0.08-66.7]	0.627
Pub or Bar ^v	1.62	[0.12-21.57]	0.715
Park	1.36	[0.02-94.29]	0.888
Amusement park	1.35	[0.05-38.3]	0.86

Note: Some regressors were automatically excluded due to multicollinearity.

- a. Working in warehouse settings – warehouse, haulage, distribution etc
- b. Working in close contact services – barbers, hairdressers, nail salons, tattoo studios and tanning salons, and any other services which require close contact
- c. Working in manufacturing and construction – textile, electronics, cars
- d. Working in hospitality – working in restaurants, food and drink outlets, lodging etc
- e. Education – working or attending
- f. Working in public transport- including buses, logistics, underground and trains
- g. Working in military – including the Navy, Army and Air Force
- h. Working in arts, or recreation – music, theatre, gyms, cinema, leisure centres
- i. Work related travel – including attending conferences, door-to-door sales, visiting clients
- j. Sports events - including football, cricket, rugby, tennis etc. matches, horse races and other sports events
- k. Personal exercise – e.g. at the gym, swimming, running clubs, casual exercise with friends
- l. Eating out – in different restaurants, cafes, pubs
- m. Charity and community – fundraising, volunteering, boot sale, corporate events, educational classes
- n. Private events – including weddings, funerals, parties, other social gatherings
- o. Mass gatherings – raves, gigs, festivals, protests and live music events
- p. Personal care - Visiting nail salons, hairdressers, barbers, tanning and tattoo studios
- q. Visiting family or friends – indoors or outdoors
- r. Travel - any international and domestic travel including trains and flights
- s. Health care - visiting healthcare services for non-COVID-19 reasons
- t. Shopping at supermarkets, local shops and pharmacies
- u. Other entertainment activities (non-specific)
- v. Going to pub or bar (entertainment, not e.g. eating out)

Table A1-3: North (incl. North East, North West, Yorkshire and Humber)

Setting/Activity	Odds Ratio	95% CI	P-value
Work and education			
Warehouse ^a	8.54	[2.29-31.83]	0.001
Military ^g	3.88	[0.19-78.03]	0.376
Healthcare	3.45	[2.16-5.51]	< 0.001
Construction ^c	2.61	[1.56-4.36]	< 0.001
Social care	2.21	[0.98-4.97]	0.056
Hospitality ^d	2.13	[1.1-4.1]	0.025
Education ^e	2.02	[1.43-2.84]	< 0.001
Transport ^f	2.02	[0.99-4.12]	0.054
Retail	1.74	[0.95-3.16]	0.072
Close contact ^b	1.64	[0.44-6.15]	0.463
Emergency services	1.35	[0.44-4.15]	0.597
Arts and recreation ^h	1.04	[0.37-2.94]	0.934
Food industry	0.45	[0.12-1.63]	0.222
Work related travel ⁱ	0.42	[0.15-1.14]	0.088
Leisure activities			

Setting/Activity	Odds Ratio	95% CI	P-value
Personal exercise ^k	0.94	[0.64-1.4]	0.767
Eating out ^l	0.69	[0.51-0.94]	0.02
Private events ⁿ	0.69	[0.16-3.04]	0.623
Sports events ^j	0.6	[0.17-2.11]	0.425
Community and charity events ^m	0.6	[0.15-2.47]	0.482
Personal care ^p	0.38	[0.18-0.81]	0.012
Visiting family ^q	0.38	[0.25-0.59]	< 0.001
Visiting healthcare ^s	0.32	[0.19-0.53]	< 0.001
Travel ^r	0.29	[0.12-0.72]	0.007
Public events, mass gatherings ^o	0.22	[0.07-0.7]	0.01
Shopping ^t	0.16	[0.13-0.2]	< 0.001
Entertainment activities	-		
Pub or Bar ^v	5.16	[0.99-27]	0.052
Other entertainment ^u	5.05	[0.64-39.58]	0.123
Visiting a zoo	1.3	[0.16-10.26]	0.807
Park	0.84	[0.02-32.64]	0.927
Social club	0.19	[0.01-5.71]	0.336

Note: Some regressors were automatically excluded due to multicollinearity.

- a. Working in warehouse settings – warehouse, haulage, distribution etc
- b. Working in close contact services – barbers, hairdressers, nail salons, tattoo studios and tanning salons, and any other services which require close contact
- c. Working in manufacturing and construction – textile, electronics, cars
- d. Working in hospitality – working in restaurants, food and drink outlets, lodging etc
- e. Education – working or attending
- f. Working in public transport- including buses, logistics, underground and trains
- g. Working in military – including the Navy, Army and Air Force
- h. Working in arts, or recreation – music, theatre, gyms, cinema, leisure centres
- i. Work related travel – including attending conferences, door-to-door sales, visiting clients
- j. Sports events - including football, cricket, rugby, tennis etc. matches, horse races and other sports events
- k. Personal exercise – e.g. at the gym, swimming, running clubs, casual exercise with friends
- l. Eating out – in different restaurants, cafes, pubs
- m. Charity and community – fundraising, volunteering, bootsale, corporate events, educational classes
- n. Private events – including weddings, funerals, parties, other social gatherings
- o. Mass gatherings – raves, gigs, festivals, protests and live music events
- p. Personal care - Visiting nail salons, hairdressers, barbers, tanning and tattoo studios
- q. Visiting family or friends – indoors or outdoors
- r. Travel - any international and domestic travel including trains and flights
- s. Health care - visiting healthcare services for non-COVID-19 reasons
- t. Shopping at supermarkets, local shops and pharmacies
- u. Other entertainment activities (non-specific)
- v. Going to pub or bar (entertainment, not e.g. eating out)

Table A1-4: South (incl South West, South East, East of England)

Setting/Activity	Odds Ratio	95% CI	P-value
Work and education			
Warehouse ^a	8.42	[0.48-148.24]	0.145
Immigration	6.95	[0.24-203]	0.26
Emergency services	4.09	[0.49-34]	0.193

Setting/Activity	Odds Ratio	95% CI	P-value
Military ^g	3.6	[0.52-24.82]	0.193
Hospitality ^d	3.12	[1.17-8.33]	0.023
Close contact ^b	2.38	[0.27-21.05]	0.437
Healthcare	2.36	[1.22-4.6]	0.011
Social care	2.29	[0.78-6.78]	0.133
Food industry	2.29	[0.29-18]	0.431
Construction ^c	1.92	[0.78-4.74]	0.157
Arts and recreation ^h	1.68	[0.32-8.74]	0.539
Education ^e	1.57	[0.96-2.59]	0.073
Transport ^f	1.14	[0.36-3.68]	0.823
Retail	0.97	[0.44-2.16]	0.944
Work related travel ⁱ	0.81	[0.16-4.15]	0.796
Leisure activities			
Community and charity events ^m	1.96	[0.05-75.9]	0.717
Eating out ^l	1.17	[0.78-1.76]	0.442
Sports events ^j	1.04	[0.28-3.89]	0.957
Personal exercise ^k	0.95	[0.55-1.63]	0.847
Travel ^f	0.91	[0.35-2.37]	0.846
Private events ⁿ	0.72	[0.08-6.16]	0.763
Personal care ^p	0.56	[0.2-1.56]	0.268
Visiting family ^q	0.38	[0.24-0.62]	< 0.001
Public events, mass gatherings ^o	0.35	[0.07-1.71]	0.195
Shopping ^t	0.17	[0.12-0.24]	< 0.001
Visiting healthcare ^s	0.14	[0.06-0.33]	< 0.001
Entertainment activities			
Visiting a zoo	12.28	[0.86-176.42]	0.065
Other entertainment ^u	7.55	[1-57.13]	0.05
Amusement park	1.92	[0.13-29.55]	0.639
Pub or Bar ^v	1.77	[0.31-10.22]	0.524
Park	0.68	[0.04-13.02]	0.801

Note: Some regressors were automatically excluded due to multicollinearity.

- a. Working in warehouse settings – warehouse, haulage, distribution etc
- b. Working in close contact services – barbers, hairdressers, nail salons, tattoo studios and tanning salons, and any other services which require close contact
- c. Working in manufacturing and construction – textile, electronics, cars
- d. Working in hospitality – working in restaurants, food and drink outlets, lodging etc
- e. Education – working or attending
- f. Working in public transport- including buses, logistics, underground and trains
- g. Working in military – including the Navy, Army and Air Force
- h. Working in arts, or recreation – music, theatre, gyms, cinema, leisure centres
- i. Work related travel – including attending conferences, door-to-door sales, visiting clients
- j. Sports events - including football, cricket, rugby, tennis etc. matches, horse races and other sports events
- k. Personal exercise – e.g. at the gym, swimming, running clubs, casual exercise with friends
- l. Eating out – in different restaurants, cafes, pubs
- m. Charity and community – fundraising, volunteering, bootsale, corporate events, educational classes
- n. Private events – including weddings, funerals, parties, other social gatherings
- o. Mass gatherings – raves, gigs, festivals, protests and live music events
- p. Personal care - Visiting nail salons, hairdressers, barbers, tanning and tattoo studios
- q. Visiting family or friends – indoors or outdoors
- r. Travel - any international and domestic travel including trains and flights
- s. Health care - visiting healthcare services for non-COVID-19 reasons
- t. Shopping at supermarkets, local shops and pharmacies

- u. *Other entertainment activities (non-specific)*
- v. *Going to pub or bar (entertainment, not e.g. eating out)*

Appendix 2.

Population Attributable Fraction

PAF is calculated using the following formula:

$$\mathbf{PAF} = \frac{\sum_{i=1}^n P_i RR_i - \sum_{i=1}^n P'_i RR_i}{\sum_{i=1}^n P_i RR_i}$$

Where:

- P_i = proportion of population at exposure level i , current exposure
- P'_i = proportion of population at exposure level i , counterfactual or ideal level of exposure
- RR (/OR) = the relative risk at exposure level i
- n = the number of exposure levels

For more information, see:

https://www.who.int/healthinfo/global_burden_disease/metrics_paf/en/

Appendix 3.

3.1 Exclusion criteria for controls and cases with excessive responses



Exclusion_Criteria_C
ontrols.csv

Appendix 4.

Table A4-1: Multivariable analysis results from the 1st iteration of the study.

Setting/Activity	Raw data			Removing excessively responding subjects*		
	Odds Ratio	95% C.I.	P-value	Odds Ratio	95% C.I.	P-value
Workplace						
Social care	5.23	(3.35 to 8.16)	<0.001	7.48	(4.47 to 12.52)	<0.001
Close contact services ^a	4.09	(1.23 to 13.63)	0.02	6.61	(1.45 to 30.15)	0.02
Health care	2.71	(1.95 to 3.76)	<0.000	3.45	(2.43 to 4.91)	<0.001
Warehouse	1.97	(0.73 to 5.33)	0.18	3.17	(1.00 to 10.06)	0.05
Hospitality	2.81	(1.76 to 4.47)	<0.001	3.16	(1.94 to 5.16)	<0.001
Food production	0.99	(0.39 to 2.48)	0.98	2.15	(0.63 to 7.36)	0.22
Transport ^b	1.17	(0.67 to 2.04)	0.59	1.88	(0.99 to 3.56)	0.05
Construction	1.28	(0.84 to 1.94)	0.26	1.48	(0.96 to 2.29)	0.08
Retail	0.99	(0.70 to 1.38)	0.93	1.47	(1.01 to 2.15)	0.05
Arts and recreation ^c	0.75	(0.34 to 1.61)	0.46	1.05	(0.45 to 2.44)	0.91
Emergency services	0.56	(0.28 to 1.10)	0.09	0.77	(0.37 to 1.60)	0.49
Immigration service	0.9	(0.03 to 23.19)	0.95	0.77	(0.03 to 20.21)	0.89
Work related travel ^d	0.51	(0.26 to 1.01)	0.05	0.63	(0.31 to 1.30)	0.21
Education ^e	0.34	(0.23 to 0.50)	<0.001	0.43	(0.29 to 0.64)	<0.001
Activities						
Entertainment ^f	2.37	(1.70 to 3.31)	<0.001	2.63	(1.86 to 3.71)	<0.001
Private events	1.5	(0.89 to 2.53)	0.13	2.1	(1.18 to 3.74)	0.01
Sports events ^g	1.06	(0.57 to 2.00)	0.85	1.23	(0.62 to 2.43)	0.55
Mass gathering	0.54	(0.28 to 1.06)	0.07	0.89	(0.44 to 1.83)	0.76
Eating out	0.78	(0.65 to 0.92)	0.004	0.8	(0.67 to 0.96)	0.02
Exercise ^h	0.62	(0.48 to 0.81)	<0.001	0.67	(0.51 to 0.87)	0.003
Charity event	0.39	(0.17 to 0.92)	0.03	0.63	(0.25 to 1.57)	0.32
Travel ⁱ	0.56	(0.44 to 0.72)	<0.001	0.59	(0.46 to 0.77)	<0.001
Health care contact ^j	0.24	(0.16 to 0.37)	<0.001	0.3	(0.19 to 0.47)	<0.001
Visiting family	0.26	(0.22 to 0.32)	<0.001	0.27	(0.22 to 0.33)	<0.001
Personal care ^k	0.2	(0.12 to 0.33)	<0.001	0.22	(0.13 to 0.37)	<0.001
Shopping ^l	0.14	(0.12 to 0.17)	<0.001	0.14	(0.12 to 0.17)	<0.001

Table A4-2: Multivariable analysis results from the 2nd iteration of the study

Setting/Activity	Odds Ratio	95% C.I.		P-value
Work and education				
Education	0.97	0.82	1.15	0.77
Arts and recreation	0.67	0.43	1.05	0.08
Emergency services	1.27	0.51	3.15	0.60
Healthcare	1.46	1.10	1.94	0.01
Social care	4.75	2.79	8.09	<0.001
Retail	0.65	0.49	0.87	0.00
Close contact	2.24	0.89	5.66	0.09
Hospitality	2.09	1.36	3.21	0.00
Construction	0.86	0.60	1.23	0.42
Warehouse	2.54	1.07	6.04	0.04
Food industry	0.37	0.19	0.73	0.00
Transport	0.72	0.45	1.14	0.16
Work related travel	0.90	0.47	1.73	0.76
Military	2.06	0.47	8.98	0.34
Immigration	0.70	0.08	6.05	0.75
Leisure activities				
Eating out	0.88	0.80	0.97	0.01
Spots events	1.08	0.63	1.84	0.78
Private events	0.80	0.41	1.55	0.51
Public events, mass gatherings	0.17	0.07	0.43	<0.001
Community and charity events	0.19	0.06	0.64	0.01
Personal exercise	1.01	0.81	1.26	0.93
Personal care	0.49	0.30	0.78	0.00
Shopping	0.32	0.29	0.35	<0.001
Travel	0.68	0.41	1.13	0.13
Visiting family	0.42	0.34	0.51	<0.001
Visiting healthcare	0.33	0.23	0.47	<0.001
Entertainment activities				
Entertainment - Pub or Bar	2.88	1.60	5.15	0.00
Entertainment - Cinema	0.08	0.02	0.29	0.00
Entertainment - Festival	0.02	0.00	0.55	0.02
Entertainment - Live music	0.86	0.11	6.90	0.89
Entertainment - Music event	0.85	0.26	2.73	0.78
Entertainment - Nightclub	0.80	0.20	3.29	0.76
Entertainment - Other	5.15	1.80	14.73	0.00
Entertainment - Park	3.18	0.76	13.30	0.11
Entertainment - Play date	0.18	0.01	3.55	0.26
Entertainment - Social club	2.80	0.74	10.64	0.13
Entertainment - Theatre	1.15	0.03	43.16	0.94
Entertainment - Day trips	4.23	0.52	34.59	0.18
Entertainment - Visiting a zoo	0.43	0.04	4.78	0.49
Entertainment - Comedy club	2.19	0.01	626.38	0.79